

# **Automatisierte Objekterkennung zur Interpretation hochauflösender Bilddaten in der Erdfernerkundung**

## **D I S S E R T A T I O N**

**zur Erlangung des akademischen Grades  
doctor rerum naturalium  
(Dr. rer. nat.)  
im Fach Informatik**

**eingereicht an der  
Mathematisch-Naturwissenschaftlichen Fakultät II  
Humboldt-Universität zu Berlin**

von

**Herrn Dipl.-Inf. Stefan Mayer**  
geboren am 15.04.1971 in Mannheim

Präsident der Humboldt-Universität zu Berlin:  
Prof. Dr. Jürgen Mlynek

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät II:  
Prof. Dr. Elmar Kulke

Gutachter:

1. Prof. Dr. Beate Meffert
2. Prof. Dr. Patrick Hostert
3. Prof. Dr. Ralf Reulke

Eingereicht am: 11. August 2003  
Tag der mündlichen Prüfung: 13. Februar 2004



## **Zusammenfassung**

Als Datengrundlage für die Erhebung von Flächennutzungsparametern, wie sie in geografischen Informationssystemen (GIS) abgelegt und verwaltet werden, dienen oft Bilddaten aus der Erdfernerkundung. Die zur Erkennung und Unterscheidung der Objekte notwendige hohe Pixelauflösung führt bei der Erfassung eines Zielgebiets wie beispielsweise einer Stadt zu enormen Datenmengen. Aus diesen Bilddaten gilt es, möglichst schnell und preiswert die für ein GIS notwendigen Informationen, wie Umrissvektoren und Objektattribute, zu extrahieren. Diese Arbeit ist ein Beitrag zur Automatisierung dieses Schritts mit besonderem Schwerpunkt auf der Gebäudeextraktion.

Datengrundlage sind hochauflösende multispektrale Orthobilder und ein digitales Oberflächenmodell (DOM) der digitalen Luftbildkamera HRSC-A bzw. HRSC-AX zum Einsatz. Deswegen werden das Aufnahmeprinzip sowie die Datenverarbeitung der HRSC überblicksartig vorgestellt.

Auf Basis dieser HRSC-Standarddatenprodukte wird ein Vorgehen zur Extraktion von Objekten entwickelt. In einer hierarchisch geordneten Abfolge an Segmentierungsschritten werden aus der Pixelinformation bedeutungstragende Einheiten extrahiert. Dieser Segmentierungsansatz lässt sich auf mehrere Objektkategorien, wie Straßen oder Ackerflächen, erweitern. So werden in der aktuellen Entwicklungsstufe neben Gebäuden auch Baumregionen detektiert.

Anhand des Oberflächenmodells werden erhöhte Regionen erkannt. Dazu wird das DOM durch Berechnung eines Terrainmodells auf Grundhöhe normiert. Für erhöhte Objekte wird die Grundhöhe aus umliegenden Grundregionen abgeleitet.

Die erhöhten Regionen werden anschließend in Bäume und Gebäude unterteilt. Dazu werden aus den Multispektraldaten Vegetationscharakteristika bestimmt und entsprechende Baumsegmente ermittelt. Die Gebäuderegionen resultieren aus einer Nachverarbeitung der verbleibenden Segmente.

Um Gebäudekomplexe in einzelne Häuser aufzuteilen, wird ein gradientenbasierter Ansatz entwickelt. Anhand der für Brandmauern typischen Gradienteninformation werden Linienhypothesen zur Unterteilung der Gebäudesegmente generiert. Diese werden schrittweise anhand geometrischer und radiometrischer Kriterien auf ihre Plausibilität überprüft.

Schließlich werden die ursprünglich aus dem DOM stammenden Konturen der Gebäudesegmente und deren Übereinstimmung mit Bildkanten eines Orthobildes betrachtet. In einem adaptiven Ansatz wird das Konturpolygon durch die Gradienteninformation an angrenzende Bildkanten angepasst. Zur Umsetzung typischer Gebäudegeometrien wie rechter Winkel oder Parallelität werden innerhalb des Adaptionprozesses entsprechende Nebenbedingungen formuliert.

Die Extraktion erhöhter Objekte wie auch deren Unterteilung in Bäume und Gebäude erfolgt mit hoher Genauigkeit, z. B. liegen die Detektionsraten bei Gebäuden über 90 %. Der neuartige Ansatz zur Unterteilung in einzelne Häuser ohne explizite Liniendetektion führt bereits in der vorgestellten Entwicklungsstufe zur Beschleunigung einer manuellen Interpretation. Die adaptive Verbesserung der Gebäudekontur führt zu gebäudetypischeren Umrissen ohne Beeinträchtigung der hohen Detektionsraten.

### **Schlagwörter:**

Objektextraktion, Hierarchischer Segmentierungsprozess, Gebäudedetektion, Adaptive Konturüberarbeitung, Hochauflösende Orthobilddaten



## **Abstract**

Remote sensing image data are often used as a basis for determining land use parameters, as they are stored and managed in geographic information systems (GIS). Covering a target area leads to an enormous amount of data due to the high pixel resolution required for recognizing and discriminating objects. To effectively derive GIS information like contour vectors or object attributes from these data, the extraction process has to be fast and cost-effective. This thesis is a contribution to the automatization of this step with a focus on building extraction.

High resolution multispectral ortho-images and a digital surface model (DSM), generated by the digital aerial camera HRSC" A or HRSC" AX, are used as data basis. Therefore, the HRSC imaging principle and data processing are summarized.

Based on these HRSC standard data products, an object extraction scheme is developed. In a hierarchically ordered sequence of segmentation steps, meaningful units are extracted from pixel information. This segmentation approach is extendable to several object categories like streets or fields. Thus, tree regions, as well as buildings are detected in the current stage of implementation.

Elevated regions are recognized using the digital surface model. For that purpose the DSM is normalized by calculating a terrain model. For elevated objects the terrain height is derived from surrounding ground regions.

Subsequently, the elevated regions are separated into trees and buildings. Determining spectral characteristics of vegetation from the multispectral data leads to corresponding tree segments. The building regions result from post-processing the remaining segments.

In order to split the building segments into single houses, a gradient based approach is developed. By means of the gradient information associated with firewalls, line hypotheses for subdividing the building segments are generated. Their plausibility is checked by gradually applying geometric and spectral criteria.

Finally, the building contours, originally derived from the DSM, and their correspondence to image edges in an ortho-image, are considered. In an adaptive approach, the contour polygon is adjusted to neighboring image edges using the gradient information. Typical building geometries like right angles or parallelism are enforced by applying corresponding constraints in the adaption process.

The extraction of elevated objects, as well as the separation into trees and buildings, is carried out with high accuracy, e. g. the building detection rates are over 90 %. In the current development stage the novel approach for separating building segments into single houses without an explicit line detection already leads to a speeding-up of a manual interpretation. The adaptive improvement of building contours leads to building typical contours without affecting the high detection rates.

### **Keywords:**

Object extraction, hierarchical segmentation process, building detection, adaptive contour improvement, high resolution ortho-images



# Danksagung

An dieser Stelle danke ich all denjenigen, die mir das Zustandekommen der Doktorarbeit ermöglicht und mich bei der Umsetzung unterstützt haben. In Ihrer Funktion als Instituts- und Abteilungsleiter danke ich Herrn Prof. G. Neukum und Herrn Dr. R. Jaumann für die strukturelle und finanzielle Unterstützung meiner Arbeit am Deutschen Zentrum für Luft- und Raumfahrt (DLR) in Berlin-Adlershof. Insbesondere danke ich Herrn F. Lehmann als Leiter der Arbeitsgruppe für Geowissenschaftliche Fernerkundung für meine Förderung und das unkomplizierte Miteinander in der Arbeitsgruppe.

Frau Prof. B. Meffert vom Lehrstuhl für Signalverarbeitung der Humboldt-Universität zu Berlin danke ich für die fordernde aber sehr angenehme Zusammenarbeit. Die Vorträge im Rahmen ihres Forschungsseminars und die anschließenden Diskussionsrunden waren immer auch für mich ein Gewinn. Ihre konstruktive Kritik insbesondere beim Überarbeiten des Manuskripts der Arbeit war mir eine wertvolle Hilfe. Das Resultat hat den Arbeitsaufwand bei der Umsetzung der Verbesserungsvorschläge voll gerechtfertigt.

Ich danke allen Mitarbeitern am DLR, die mich durch fachliche Gespräche gefordert und durch gemeinsame Aktivitäten auch außerhalb des DLR angenehm „abgelenkt“ haben. Dem HRSC-Team danke ich für die „hautnahen“ Einblicke in die Sensortechnik und die Bilddatenverarbeitung. Besonderer Dank gilt Herrn F. Scholten, der trotz chronischer Überlastung stets für die Diskussion photogrammetrischer Fragestellungen und für die Anfertigung von Sonderdatenprodukten ansprechbar war.

Den jetzigen und ehemaligen Mitarbeitern der Arbeitsgruppe für Geowissenschaftliche Fernerkundung am DLR danke ich für das konstruktive und freundliche Miteinander. Ganz besonders möchte ich mich bei Herrn T. Bucher und bei Herrn S. Pless für den intensiven Austausch bedanken. Herrn T. Bucher danke ich ausdrücklich für das Gegenlesen dieser Arbeit, gerade in einer für ihn sehr fordernden Zeit.

Meiner Partnerin Daniela danke ich für ihre Geduld und ihre aufmunternden Worte. Im Lauf der Arbeit durchwanderte sie einige Höhen und Tiefen mit mir. Ihre Bereitschaft, mit mir über meine fachlichen Probleme zu diskutieren, haben mich dazu veranlasst, manche Fragestellungen aus einer anderen Perspektive zu betrachten.

Dadurch, dass mir meine Mutter und mein verstorbener Vater eine entsprechende Ausbildung ermöglicht haben, wurden die Grundsteine für diese Arbeit gelegt. Ein herzliches Dankeschön dafür.

Bei meinem Freundeskreis und meiner Familie entschuldige ich mich für die mit der Arbeit verbundenen Einschränkungen. Für die wohltuenden Abwechslungen im Lauf der Arbeit bedanke ich mich sehr.





# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>1</b>
1.1	Einführung in die Interpretation von Bilddaten in der Erdfernerkundung . . . . .	1
1.2	Ausgangsdaten . . . . .	6
1.3	Literaturüberblick . . . . .	7
1.3.1	Grundlegende Konzepte der Gebäudeextraktion . . . . .	7
1.3.2	Modellfreie Ansätze zur Gebäudeextraktion . . . . .	10
1.4	Ziel und Aufbau der Arbeit . . . . .	12
1.4.1	Zielsetzung . . . . .	12
1.4.2	Vorgehen zur Objektextraktion . . . . .	13
1.4.3	Struktur der Arbeit . . . . .	17
1.5	Implementierungsaspekte . . . . .	19
1.5.1	Die Programmierungsumgebung . . . . .	19
1.5.2	Anforderungen hochauflösender Daten an die Implementierung von Bildver- arbeitungsroutinen . . . . .	20
1.5.3	Bildkanten als Objektgrenzen . . . . .	22
<b>2</b>	<b>Der Multispektralscanner HRSC-A</b>	<b>25</b>
2.1	Aufbau und Eigenschaften . . . . .	25
2.2	Das Aufnahmeprinzip . . . . .	27
2.3	Die HRSC-Datenverarbeitung . . . . .	31
2.3.1	Generierung der Bildstreifen . . . . .	31
2.3.2	Radiometrische Korrektur . . . . .	32
2.3.3	Geometrische Korrektur . . . . .	33
2.3.4	Generierung eines Oberflächenmodells . . . . .	35
2.3.5	Orthobild-Generierung . . . . .	36
2.3.6	Mosaik-Erstellung . . . . .	36
2.4	Geometrische Qualität . . . . .	37
2.5	Die automatisierte Interpretation von HRSC-Daten . . . . .	37
2.5.1	Radiometrie . . . . .	37
2.5.2	Oberflächenmodell . . . . .	40
<b>3</b>	<b>Extraktion erhöhter Objekte aus einem Oberflächenmodell</b>	<b>45</b>
3.1	Einführung und Literaturüberblick . . . . .	45
3.2	Interpolation der Grundhöhe . . . . .	47
3.3	Pyramidale Verfeinerung der Geländehöhe . . . . .	51
3.4	Evaluation der DTM-Berechnung . . . . .	55

3.4.1	Vergleich der beiden Ansätze . . . . .	55
3.4.2	Qualitative und quantitative Bewertung . . . . .	57
3.5	Objektextraktion . . . . .	61
<b>4</b>	<b>Segmentierung erhöhter Objektregionen</b>	<b>65</b>
4.1	Einführung und Literaturüberblick . . . . .	65
4.2	Vegetationssegmentierung . . . . .	68
4.2.1	Der Vegetationsindex NDVI . . . . .	69
4.2.2	Farbtonsegmentierung der Vegetation . . . . .	75
4.2.3	Kombination von NDVI und Farbtonsegmentierung zur Identifikation von Vegetation . . . . .	77
4.2.4	Vereinfachtes Vorgehen für HRSC-AX Daten . . . . .	79
4.3	Nachverarbeitung der Gebäudesegmente . . . . .	80
4.3.1	Schattensegmentierung . . . . .	80
4.3.2	Morphologisches Öffnen . . . . .	83
4.4	Evaluation . . . . .	84
4.4.1	Trennbarkeit von Bäumen und Gebäuden bei HRSC-A und HRSC-AX . . .	86
4.4.2	Baumdetektion . . . . .	90
4.4.3	Gebäudedetektion . . . . .	92
<b>5</b>	<b>Segmentierung in einzelne Objekte</b>	<b>97</b>
5.1	Einführung und Literaturüberblick . . . . .	98
5.2	Begriffsdefinition und Vorverarbeitungsschritte . . . . .	100
5.3	Trennlinien durch ausgewählte Skelettpixel . . . . .	101
5.4	Lage der Linien im Gebäudesegment . . . . .	104
5.4.1	Ausschluss von Linien ohne beidseitige Segmentfortsetzung . . . . .	104
5.4.2	Verkürzung der Linienenden . . . . .	105
5.4.3	Ausschluss von Linien nahe des Segmentrands . . . . .	107
5.5	Kriterien zur Beurteilung von Trennlinien . . . . .	107
5.5.1	Mittelung der Gradientenstärke . . . . .	107
5.5.2	Differenz der beidseitigen Grauwerte . . . . .	110
5.6	Bewertung und Auswahl geeigneter Trennlinien . . . . .	111
5.6.1	Entfernen benachbarter Linien . . . . .	112
5.6.2	Entfernen von Linien, die zu zu kleinen Segmenten führen würden . . . . .	112
5.7	Realisierung der Trennlinien . . . . .	113
5.8	Evaluation der Gebäudeunterteilung . . . . .	116
5.8.1	Qualitative Beurteilung . . . . .	116
5.8.2	Genauigkeit . . . . .	119
5.8.3	Rechenzeit . . . . .	122
<b>6</b>	<b>Nachbearbeitung der Gebäudekonturen</b>	<b>125</b>
6.1	Einführung und Literaturüberblick . . . . .	126
6.2	Vorverarbeitung der Bilddaten . . . . .	129
6.2.1	Polygonvereinfachung . . . . .	129
6.2.2	Lokale Gradientenstärke . . . . .	130

6.3	Adaption an Bildkanten . . . . .	131
6.3.1	Ermittlung der Polygongüte . . . . .	132
6.3.2	Optimierung der Konturgüte . . . . .	134
6.4	Berücksichtigung gebäudetypischer Geometrien . . . . .	137
6.4.1	Einführen gebäudespezifischer Nebenbedingungen . . . . .	138
6.4.2	Umsetzung der Nebenbedingungen . . . . .	141
6.5	Dynamische Verwaltung der Stützstellen . . . . .	145
6.5.1	Löschen von Vertices . . . . .	145
6.5.2	Ausprägen rechter Winkel . . . . .	145
6.5.3	Einfügen neuer Vertices . . . . .	146
6.5.4	Verschieben von Vertices . . . . .	147
6.6	Ablauf der Konturüberarbeitung . . . . .	147
6.7	Evaluation der Konturüberarbeitung . . . . .	152
6.7.1	Qualitative Beurteilung . . . . .	155
6.7.2	Genauigkeit . . . . .	159
6.7.3	Rechenzeit . . . . .	164
<b>7</b>	<b>Zusammenfassung und Ausblick</b>	<b>167</b>
7.1	Zusammenfassung der Arbeit . . . . .	167
7.2	Ausblick . . . . .	171
	<b>Symbol- und Abkürzungsverzeichnis</b>	<b>177</b>
	<b>Literaturverzeichnis</b>	<b>181</b>



# Abbildungsverzeichnis

1.1	Bildausschnitt aus der Berliner Innenstadt überlagert mit Konturlinien der ALK Gebäudeebene . . . . .	3
1.2	Arbeitsplätze zur Interpretation von Stereo-Luftbildern . . . . .	4
1.3	Schematischer Überblick des hierarchischen Vorgehens zur Objektextraktion mittels multispektraler Orthobilder und eines DOM . . . . .	15
1.4	Die ENVI Arbeitsumgebung . . . . .	20
1.5	Bildbeispiel zur Kachelung großer Bilddaten . . . . .	21
2.1	Die zwei Generationen der HRSC-A . . . . .	26
2.2	Das Pushbroom-Scanprinzip der HRSC . . . . .	29
2.3	Darstellung der HRSC-Datenverarbeitung . . . . .	31
2.4	Bildbeispiele zur photogrammetrischen Verarbeitung . . . . .	34
2.5	Schematische Darstellung zur Blickwinkelproblematik . . . . .	39
2.6	Bildbeispiel zur Interpolation des blickwinkelbedingt nicht erfassten Grundbereichs in der Ortho-Korrektur . . . . .	40
2.7	Veranschaulichung des DOM-Interpolationsfehlers . . . . .	41
2.8	Veranschaulichung des Kanalversatzes nach der Ortho-Korrektur . . . . .	42
3.1	Exemplarische Veranschaulichung des Vorgehens zur Detektion erhöhter Punkte für die Grundhöheninterpolation an einem Höhenprofil . . . . .	49
3.2	Schematische Darstellung des Ablaufs der Grundhöheninterpolation anhand benachbarter Grundpunkte . . . . .	49
3.3	Auswirkung der Größe des Nachbarschaftsfensters und der Höschwelle auf die Detektion erhöhter Regionen . . . . .	50
3.4	Schematische Darstellung des Ablaufs der pyramidalen Verfeinerung der Grundhöhe . . . . .	52
3.5	Fallbeispiele zur histogrammbasierten Bestimmung der Grundhöhe . . . . .	54
3.6	Exemplarische Veranschaulichung des Vorgehens zur pyramidalen Verfeinerung der Grundhöhe an einem Höhenprofil . . . . .	55
3.7	Testdatensatz zur Beurteilung der Methoden zur DOM-Normierung . . . . .	58
3.8	Terrainmodelle, gewonnen durch Grundinterpolation und pyramidale Verfeinerung der Grundhöhe . . . . .	59
3.9	Gegenüberstellung eines DOM-Höhenprofils mit Profilen von DTMs aus Grundinterpolation und pyramidalen Verfeinerung der Grundhöhe . . . . .	60
3.10	Durch Grundinterpolation normiertes Oberflächenmodell . . . . .	62
4.1	Schematische Darstellung zur Segmentierung in Vegetation und Gebäude . . . . .	67
4.2	Schematische Darstellung zur Vegetationssegmentierung . . . . .	68

4.3	Gegenüberstellung des Transmissionsbereichs von HRSC-Filtern und der Reflektanz ausgewählter Materialien . . . . .	71
4.4	Schematische Darstellung der Bandbreiten des roten und panchromatischen HRSC-A Kanals sowie eines daraus abgeleiteten künstlichen Rotkanals . . . . .	73
4.5	Farbkomposite aus HRSC-A Daten und daraus abgeleitetem künstlichen Rotkanal . .	73
4.6	HRSC-A Farbkomposit und zugehörige Grauwertdarstellung des NDVI . . . . .	74
4.7	Schematische Darstellung des HSV-Farbraums . . . . .	75
4.8	HRSC-A Farbkomposit und zugehöriger Nadirausschnitt überlagert mit dem Resultat der Vegetationssegmentierung . . . . .	78
4.9	Schematische Darstellung zur Nachverarbeitung der Gebäudesegmente . . . . .	80
4.10	HRSC-A Farbkomposit und zugehöriger Nadirausschnitt überlagert mit dem Resultat der Schattensegmentierung . . . . .	83
4.11	Bildbeispiel zur morphologischen Nachverarbeitung der Gebäudesegmente . . . . .	84
4.12	HRSC-A Farbkomposit und zugehöriger Nadirausschnitt überlagert mit dem Resultat der Gebäudesegmentierung . . . . .	85
4.13	Resultat der Gebäude- und Vegetationssegmentierung . . . . .	86
4.14	Relative Pixelhäufigkeiten in Abhängigkeit vom NDVI bei HRSC-A und HRSC-AX Daten . . . . .	87
4.15	Relative Häufigkeiten von Regionen in Abhängigkeit vom zugeordneten mittleren NDVI-Wert bei HRSC-A und HRSC-AX Daten . . . . .	88
4.16	Relative Pixelhäufigkeiten in Abhängigkeit vom Farbton und NDVI bei Daten der HRSC-A . . . . .	89
4.17	Gegenüberstellung von Baumsegmentierung und manueller Vorgabe . . . . .	91
4.18	Fehlzuordnungen in der Baumsegmentierung . . . . .	93
5.1	Brandmauern am Übergang zweier aneinandergrenzender Häuser . . . . .	98
5.2	Schematische Darstellung zur Aufteilung der Gebäudesegmente in kleinere Einheiten	99
5.3	Schematische Darstellung zur Entwicklung eines Trennlinienkandidaten . . . . .	102
5.4	Bestimmung von $\hat{\varphi}$ durch Betrachtung des Histogramms der Gradientenorientierung .	103
5.5	Schematische Darstellung des Zusammenhangs von lokaler Skelettpixelausrichtung und Hauptgradientenorientierung . . . . .	103
5.6	Schematische Darstellung zum Ausschluss von Linien, für die es keine beidseitige Fortsetzung im Gebäudesegment gibt . . . . .	105
5.7	Schematische Darstellung zur Verkürzung der Trennlinien an deren Enden . . . . .	106
5.8	Schematische Darstellung zum Ausschluss von Linien, die länger entlang des Segmentrands verlaufen . . . . .	107
5.9	Veranschaulichung zur Beurteilung der Verteilung passender Kantenpixel entlang der Trennlinie . . . . .	109
5.10	Schematische Darstellung zur Bildung der beidseitigen Grauwertdifferenz von Trennlinien . . . . .	111
5.11	Schematische Darstellung zur Bestimmung von benachbarten Linien ähnlicher Orientierung . . . . .	113
5.12	Schematische Darstellung zur Ermittlung des besten Pfades zur Kombination „innerer“ Trennlinien . . . . .	115
5.13	Zwischenresultate bei der Unterteilung eines Gebäudkomplexes in kleinere Einheiten	117

5.14	Endergebnis bei der Unterteilung eines Gebäudkomplexes in kleinere Einheiten . . .	118
5.15	Auswertung der Unterteilung von Gebäudkomplexen an einem Referenzdatensatz . .	119
5.16	Veranschaulichung des Vorgehens zur Bewertung des Trennlinienalgorithmus . . . .	120
6.1	Vereinfachte Darstellung zur Nachbearbeitung der Gebäudekonturen . . . . .	128
6.2	Bildbeispiel zur Polygonvereinfachung und Ermittlung der lokalen Gradientenstärke $ \nabla ^{(loc)}$ . . . . .	131
6.3	Schematische Darstellung zur Nomenklatur der geometrischen Beziehungen inner- halb eines Polygons . . . . .	133
6.4	Schematische Darstellung zur Adaption der Polygonstützpunkte auf dem diskreten Bildraster . . . . .	134
6.5	Schematische Darstellung zur Optimierung der Polygonstützpunkte nach dem Prin- zip der Dynamischen Programmierung . . . . .	135
6.6	Schematische Darstellung zur lokalen Optimierung der Polygonstützpunkte . . . .	137
6.7	Schematische Darstellung zur Gruppierung in aufeinander folgende Polygonvertices mit nahezu rechtem Winkel . . . . .	141
6.8	Schematische Darstellung zum Einfügen neuer Polygonvertices . . . . .	146
6.9	Schematische Darstellung zum zufälligen Verschieben schwacher Polygonvertices . .	147
6.10	Schematische Darstellung des Ablaufs zur Überarbeitung der Gebäudekonturen . . .	148
6.11	Konturüberarbeitung an einem Bildbeispiel — erste Optimierungsschleife (Teil 1) . .	153
6.12	Konturüberarbeitung an einem Bildbeispiel — erste Optimierungsschleife (Teil 2) . .	154
6.13	Konturüberarbeitung an einem Bildbeispiel — zweite Optimierungsschleife . . . .	155
6.14	Künstliche Verschlechterung der Gebäudekonturvorgaben und anschließende Nach- bearbeitung . . . . .	156
6.15	Nachbearbeitung der aus der Vorgabe abgeleiteten Umrisspolygone . . . . .	157
6.16	Nachbearbeitung der automatisiert ermittelten Umrisspolygone . . . . .	159
6.17	Gegenüberstellung der hausweise berechneten Detektionsraten vor und nach der Kon- turüberarbeitung . . . . .	162
6.18	Detektionsraten in Abhängigkeit von der Komplexität der Konturlinie . . . . .	162
7.1	Bildbeispiel zum Ablauf der Objektextraktion . . . . .	171





# Tabellenverzeichnis

2.1	Technische Parameter unterschiedlicher HRSC-Modelle . . . . .	26
2.2	Spektrale Empfindlichkeit und Blickwinkel der Kanäle von unterschiedlichen HRSC-Modellen . . . . .	28
3.1	Gegenüberstellung von Vor- und Nachteilen der beiden Verfahren zur DOM-Normierung . . . . .	56
3.2	Genauigkeitsangaben für die beiden vorgestellten Methoden zur DOM-Normierung .	60
3.3	Laufzeitangaben für die beiden vorgestellten Methoden zur Ableitung eines DTM aus einem DOM . . . . .	61
4.1	Abschätzung des minimalen Klassifikationsfehlers bei der Trennung von Bäumen und Gebäuden für Daten der HRSC-A und HRSC-AX . . . . .	88
4.2	Genauigkeitsangaben für die Baumdetektion . . . . .	92
4.3	Genauigkeitsangaben für die Gebäudedetektion . . . . .	94
5.1	Genauigkeitsangaben für die Unterteilung der Gebäudesegmente . . . . .	121
5.2	Laufzeitangaben für das Unterteilen eines Gebäudesegments . . . . .	122
6.1	Genauigkeitsangaben für die beschränkte Adaption von künstlich verschlechterten Gebäudeumrissen an Bildkanten . . . . .	160
6.2	Genauigkeitsangaben für die beschränkte Konturoptimierung bei Gebäudekomplexen, die aus vorhergehenden Arbeitsschritten resultieren . . . . .	163
6.3	Laufzeitangaben für die Berechnung der lokalen Gradientenstärke . . . . .	165
6.4	Laufzeitangaben für die beschränkte Konturoptimierung . . . . .	165



# Algorithmenverzeichnis

1	Berechnung eines DTM durch Grundhöheninterpolation . . . . .	48
2	Berechnung eines DTM durch pyramidale Verfeinerung der Geländehöhe . . . . .	53
3	Farbraumkonvertierung von RGB zu HSV . . . . .	76
4	Berechnung einer Vegetationssegmentierung mit multispektralen HRSC-A Daten . .	77
5	Vorgehen zur Realisierung der Trennlinienkandidaten . . . . .	114
6	Konturüberarbeitung, Kommandierung der beiden Optimierungsschleifen . . . . .	149
7	Konturüberarbeitung, erste Optimierungsschleife . . . . .	150
8	Konturüberarbeitung, zweite Optimierungsschleife . . . . .	151



# 1 Einleitung

In diesem einleitenden Kapitel wird der Leser in die Thematik der Arbeit eingeführt (Abschnitt 1.1) und mit der Datengrundlage der Arbeit (Abschnitt 1.2) vertraut gemacht. Um diese Arbeit im Kontext ähnlicher Ansätze zur Objektextraktion einordnen zu können, wird in Abschnitt 1.3 ein Überblick über bestehende Forschungsansätze gegeben. In Abschnitt 1.4 wird dann die Zielstellung dieser Arbeit formuliert und die entwickelte Methodik zur Gebäudeextraktion skizziert. Zum Ende dieses Kapitels wird in Abschnitt 1.5 auf Aspekte eingegangen, die Einfluss auf die Implementierung der Methoden haben.

## 1.1 Einführung in die Interpretation von Bilddaten in der Erdfernerkundung

Der französische Luftschiffer und Fotograf Gaspard-Félix Nadar läutete 1858 mit der Erstellung der ersten Luftaufnahme von Paris die Ära der Luftbildfotografie ein. Wurden die Luftbilder in den Pionierjahren hauptsächlich von Luftschiffen und Ballons aus aufgenommen, so dienen heute vor allem Flugzeuge und Satelliten als Trägersysteme. Neben Einblicken in schwer zugängliche Gebiete ermöglicht oder vereinfacht der Wechsel des Blickpunkts über die Erdoberfläche das Erkennen großflächiger Zusammenhänge sowie die Erfassung zusätzlicher Merkmale, die beispielsweise geologisch, archäologisch, forstwirtschaftlich oder kartografisch relevant sind. In eben diesen Sparten ist die Erdfernerkundung zu einem unersetzlichen Standbein geworden.

Während bei Luftbildaufnahmen das Grundprinzip der Fotografie, die Belichtung eines lichtempfindlichen Materials, über die Jahre hinweg unverändert geblieben ist, hat sich der Aufnahme- und Auswerteprozess kontinuierlich entwickelt. So ermöglichen sogenannte Reihenmesskammern die flächendeckende Erfassung eines Zielgebiets mit einem festen Überlappungsanteil der Bildstreifen. Die Auswertung der Bilddaten erfolgt heutzutage großteils rechnergestützt. Dazu werden die Aufnahmen mit hochauflösenden Luftbildscannern abgetastet und in einem digitalen Bildformat abgelegt.

Die regelmäßige Erfassung großer Zielgebiete wurde durch den Einzug der Halbleitertechnologie in die satellitengestützte Fernerkundung möglich. Mit CCD-Zeilensensoren kann die Erdoberfläche kontinuierlich abgetastet werden. Die digitalen Bilddaten werden dann per Funk zu einer Bodenstation übertragen. Die Entwicklungen im Bereich der satellitengestützten Positions- und Lagebestim-

mung (D-GPS/INS) ermöglichen den Einsatz der Zeilensensorik nun auch in Luftbildbefliegungen. Die digitalen Sensorsysteme (u. a. die HRSC, „High Resolution Stereo Camera“) haben sich im Lauf der letzten Jahre auf dem Luftbildmarkt etabliert; eine Ablösung der immer noch stark dominierenden Luftbildfotografie wird angestrebt. Die grundsätzlichen Unterschiede liegen in der Sensorik und der Datenaufzeichnung. Durch die Georeferenzierung der Bildpunkte bereits in der Aufnahmephase wird eine sehr hohe räumliche Genauigkeit der Datenprodukte erzielt.

Um neue Bilddaten bzw. daraus abgeleitete Informationen vorhandenen Datenbeständen gegenüberstellen zu können, werden die Daten mit geografischen Koordinaten versehen („Geokodierung“). Die rechnergestützte Verwaltung, Verknüpfung und Auswertung von geokodierter Information erfolgt zunehmend über speziell darauf ausgerichtete Datenbanksysteme, die sogenannten Geografischen Informationssysteme (GIS). Im Kontext der städtischen Infrastruktur verwalten beispielsweise die Energieversorger in einem solchen GIS ihr Strom- oder Gasnetz. Mobilfunkunternehmen sind vor allem an einer Gebäudedatenbank interessiert, um die Standorte für ihre Sendemasten zu optimieren. Die Erstellung von Navigationssystemen basiert auf Geodaten. Bei der Stadt- oder Landschaftsplanung müssen viele unterschiedliche bebauungs- und umweltrelevante Aspekte berücksichtigt werden, die in einem GIS vereint und je nach Aufgabenstellung kombiniert werden können. Ferner verwalten Ämter und Behörden (z. B. Kataster- und Vermessungsamt) amtliche Geobasisdaten.

Als Beispiel amtlicher Kartenwerke sei das Liegenschaftskataster genannt, das Verzeichnis der Flurstücke für das Grundbuch. Der Gesetzgeber verlangt, dass das Liegenschaftskataster die Grundlage für raumbezogene Entscheidungen und Maßnahmen staatlicher und kommunaler Stellen bildet (DERKSEN ET AL. 2001). Schwerpunkte sind der Umwelt-, Natur- und Landschaftsschutz, die Raumordnung sowie die Bauleitung und Statistik. Per gesetzlicher Vorgabe (z. B. Vermessungs- und Liegenschaftsgesetz Brandenburg, § 1, Abs. 5) wird die Überführung der analogen Liegenschaftskarte in ein GIS, die Automatisierte Liegenschaftskarte, ALK, gefordert. Es soll ein öffentliches raumbezogenes Basisinformationssystem geschaffen werden, das die Liegenschaften nicht nur beschreibt, sondern auch darstellt. Mit der ALK soll von staatlicher Seite aus gewährleistet werden, dass allen interessierten Nutzern geografische Basisdaten *flächendeckend, aktuell und kostengünstig* zu Verfügung stehen.

Der flächendeckende Aufbau der ALK ist beispielsweise in Brandenburg aufgrund der unzureichenden analogen Datengrundlagen kurz- bis mittelfristig nicht in hoher geometrischer Genauigkeit und inhaltlicher Tiefe zu bewerkstelligen (MASUR 1997). Deshalb wurde eine ALK-Vorstufe eingeführt, deren Genauigkeit mindestens der herkömmlichen Liegenschaftskarte entsprechen muss und die bei jeder sich bietenden Gelegenheit zu verbessern ist. Ziel ist eine Überführung der ALK-Vorstufe zur Endstufe.

Für ein GIS wird die Geoinformation in unterschiedliche Kategorien aufgeteilt und in entsprechenden Informationsebenen des GIS abgelegt. Bei der ALK hat sich dafür die ursprüngliche Bezeichnung „Folie“ gehalten: Analoge Liegenschaftsdaten werden kategorieweise auf großformatige transparente Folien gezeichnet bzw. gedruckt, die je nach Bedarf miteinander überlagert werden. Ab-



**Abbildung 1.1:** Bildausschnitt aus der Berliner Innenstadt überlagert mit gelb dargestellten Konturlinien der ALK Gebäudeebene.

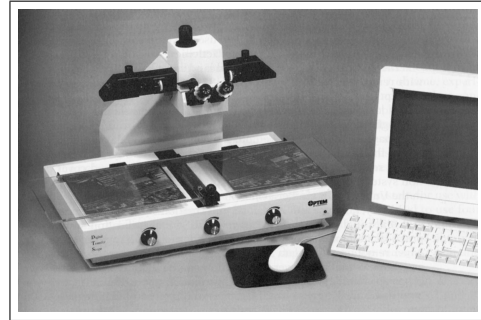
Abbildung 1.1 zeigt die Überlagerung eines HRSC-Orthobildes mit Gebäudeumrissen, die in der ALK Berlin in der Folie 11, der Gebäudeebene, abgelegt sind.

Werden Bilddaten als Basis für ein Geoinformationssystem verwendet, so gilt es, die für das GIS nötige Information aus den Bildern zu extrahieren. Auf Basis einer Luftaufnahme lassen sich die im Zentrum dieser Arbeit stehenden Umriss von Gebäuden ermitteln, aber beispielsweise auch Straßen- und Flußläufe, Gleise und Positionen von Bäumen oder Ampeln. Dies geschieht noch überwiegend manuell durch Interpretatoren mit entsprechender Expertise. Die betreffenden Polygone, Linien oder Punkte werden nachgezogen — früher mit Stift und Folie, heute am Rechner mit der Maus oder einem Digitalisierstift. Abbildung 1.2 zeigt beispielhaft Instrumente und Arbeitsplatz zur Interpretation von Luftbildern.

Gerade bei amtlichen Kartenbeständen spielt neben der Erhebung der Daten der Abgleich von neuen und vorhandenen Daten eine entscheidende Rolle. Durch vorgeschriebene regelmäßige Befliegungen soll ein aktueller Datenbestand gewährleistet werden. Die Nutzung von Bilddaten mit hoher geometrischer Genauigkeit minimiert den Aufwand für den Abgleich der Datensätze. Die Genauigkeiten der neuen digitalen Sensoren erfüllen in vielen Fällen (s. Abschnitt 2.4) die Vorgaben für amtliche Kartenprodukte und eignen sich somit auch als Referenzdatensatz.



(a) Luftbilder und Stereobrille



(b) Arbeitsplatz zur Stereointerpretation

**Abbildung 1.2:** Arbeitsplätze zur Interpretation von Stereo-Luftbildern (aus LILLESAND UND KIEFER 2000).

Die Erstellung bzw. Aktualisierung von Geoinformationssystemen führt also zu einem hohen *Bedarf an interpretierter Geoinformation*. Um die damit verbundenen Kosten möglichst gering zu halten, ist eine *großflächige Erfassung* der Geodaten und eine *hochgradig automatisierte Vorgehensweise* zur Bereitstellung der Information unerlässlich.

Für eine großflächige Abdeckung von Projektgebieten eignet sich die luftgestützte Erhebung von Bilddaten. Prinzipiell gilt: Je höher der Bildflug, desto größer ist das von der Aufnahme abgedeckte Gebiet und desto geringer die räumliche Auflösung eines Bildpunkts. Das Verhältnis von Objektgröße zu Bildauflösung gibt Aufschluss darüber, ob die interessierenden Objekte als solche noch in den Bilddaten zu erkennen sind. MULLER ET AL. (1997) diskutieren in diesem Zusammenhang die Auswirkung der räumlichen Auflösung auf eine automatisierte Gebäudeextraktion. KONECNY UND SCHIEWE (1996) geben eine Abschätzung der minimalen Grundpixelgröße zur visuellen Erkennbarkeit unterschiedlicher Objekte bei satellitengestützt gewonnenen Bilddaten wieder. So werden zur Erkennung von Häusern eine Pixelgröße von 2 m, von kleineren Straßen 5 m und von Gebäudeblocks oder großen Straßen 10 m angegeben. Bei höheren Auflösungen werden die Textur bzw. Komponenten (bei Gebäuden z. B. Kamine, Antennen oder Dachfenster) des betrachteten Objekts sichtbar. Dadurch kann die Interpretation der Bilddaten erheblich erleichtert werden. Beispielsweise nutzt MAY (2002) Fahrbahnmarkierungen, um Straßennetze zu detektieren. Die in dieser Arbeit verwendeten Standarddatenprodukte der HRSC-A weisen räumliche Auflösungen von 10 cm bis zu einem Meter auf. Großflächige Projektgebiete in einer Auflösung dieser Größenordnung zu erfassen, führt zu enormen Datenmengen, die es handzuhaben und letztendlich zu interpretieren gilt (s. dazu auch Abschnitt 1.5.2).

Bei der Erstellung der ALK wird explizit spezifiziert, dass dies auf Grundlage photogrammetrischer Auswertungen von Befliegungen erfolgen soll. Orthophotos sollen als wichtiger Baustein zur Bestimmung von Passpunkten im Hinblick auf eine Verbesserung der ALK dienen (DERKSEN ET AL. 2001). Für eine einführende Beschreibung des Datenprodukts „Orthobild“ sei auf den nach-



folgenden Abschnitt 1.2 verwiesen.

Zur Generierung von Orthobildern bietet digitale Sensorik eine nahezu vollautomatisierte Verarbeitungskette (für Daten der HRSC-A wird diese in Kapitel 2 skizziert). Das bei Fotos nötige Scannen des Filmmaterials entfällt durch die digitale Aufzeichnung.

Während der Prozess der Bilderzeugung also bereits stark automatisiert ist, erfolgt die Interpretation der Bilddaten noch überwiegend manuell durch entsprechende Experten. Zwischen der Erzeugung der Bilddaten und der rechnergestützten Verarbeitung von Geodaten in einem GIS klafft somit eine *Automatisierungslücke*. Die Experten erreichen zwar bei der Interpretation für Laien unerreichbare Geschwindigkeiten. Eine Bearbeitung großer Flächen wie ganzer Stadtgebiete erfordert dennoch einen hohen zeitlichen und somit auch finanziellen Aufwand. Nach GREGER (1996) stellen digitale Geodaten die wichtigste und gleichzeitig kostenintensivste Voraussetzung für ein Geoinformationssystem dar.

Diese Arbeit liefert in Form mehrerer Module zur Objektextraktion und Konturbearbeitung einen Beitrag zur automatisierten Interpretation hochauflösender luftgestützt gewonnener Bilddaten und somit zum Schließen der zuvor aufgezeigten „Automatisierungslücke“. Kern der Arbeit ist die *Extraktion von Gebäudekonturen* aus den nachfolgend in Abschnitt 1.2 beschriebenen Bilddaten. Die geokodierten Umrisse einzelner Häuser stellen die Basis von Geoinformationssystemen für den urbanen Bereich dar. Die Gebäudeeinheiten werden im GIS anwendungsspezifisch mit Attributen versehen. Einige der Attribute, wie Gebäudehöhen oder die Grundfläche sind direkte Resultate des vorgestellten Vorgehens.

Neben dem gerade motivierten GIS-Einsatz der in dieser Arbeit extrahierten Gebäudeinformation seien nachfolgend weitere Anwendungsfelder genannt.

In zunehmenden Maße werden Geodaten für *Visualisierungszwecke* eingesetzt. Die Anwender sollen sich beispielsweise zu Navigationszwecken oder zur touristischen Information in möglichst realistischen zwei- oder dreidimensionalen virtuellen Welten bewegen können. Damit ein nutzergesteuertes „Durchwandern“ dieser künstlichen Welten in Echtzeit möglich ist, werden die Objekte als Polygonmodelle repräsentiert. Die in dieser Arbeit ermittelten Gebäudeumrisse sind dafür ideal geeignet. Zusammen mit den Ausgangsdaten ergibt sich eine 3-D-Darstellung mit nicht nur realistisch wirkender, sondern real vorhandener Oberflächentextur.

Die genau bestimmten Umrisse von Gebäuden können ferner zur Feinkorrektur in die photogrammetrische Datenverarbeitung zurückfließen. Eine *Verbesserung der Ausgangsdaten* gerade an ihren Schwachstellen (s. folgender Abschnitt und Kapitel 2) wird so ermöglicht.

Diese Arbeit versteht sich außerdem als Vorstufe für eine *objektbasierte Bildinterpretation*. Sind die Objektumrisse einmal ermittelt, lassen sich objektspezifische Merkmale aus dem Objektsegment ableiten. Auf ihrer Basis kann das Objekt in verschiedene Kategorien klassifiziert werden, ein Baumsegment also in Baumarten wie Platane, Linde oder Kastanie. Bei geometrisch geringer aufgelösten Bildern wird dazu meist pixelweise vorgegangen. Falls die Objektabmessung die Bodenabdeckung eines Pixels um ein Mehrfaches übersteigt, ist dieses Vorgehen nicht mehr angebracht. Bedingt durch

das Zusammenspiel der Komponenten eines Objekts (z. B. Äste, Blätter) mit Licht und Schatten variieren die Reflektanzwerte des Objekts zu stark. Eine Erfassung dieser objekttypischen Textur als Klassifikationsmerkmal ist nur dann präzise, wenn möglichst keine objektfremden Pixel hinzugezogen werden. Dies wiederum ist nur bei Vorliegen der Objektgrenzen möglich („objektbasiert“).

### 1.2 Ausgangsdaten

Datengrundlage dieser Arbeit sind Standarddatenprodukte des luftgestützten Erdfernerkundungssensors HRSC-A. Zum einen sind dies *orthokorrigierte Multispektraldaten* mit einer typischen geometrischen Auflösung von 10 cm bis zu einem Meter. Fünf Spektralkanäle erfassen den sichtbaren und nah-infraroten Bereich des elektromagnetischen Spektrums. In der Erdfernerkundung werden multispektrale Daten beispielsweise erfolgreich zur Detektion von Vegetationsflächen eingesetzt.

Resultat der Orthokorrektur sind künstlich erzeugte Bildprodukte, in denen für jeden Bildpunkt die Parallelperspektive simuliert ist. Die durch die Perspektive und die Beschaffenheit des Geländes hervorgerufenen Verzerrungen sind beseitigt. Die Orthoperspektive kann die Objektextraktion sehr erleichtern: Dadurch, dass alle Objekte senkrecht von oben betrachtet werden, entstehen Bildkanten (s. Abschnitt 1.5.3) nur durch einen Helligkeitswechsel mit horizontaler Ausdehnung. Auch die Pixelgröße ist in einem Orthobild für alle Pixel gleich. Eine Umrechnung von Maßen im Bildraum in den Objektraum, also z. B. eine Längenkonvertierung von Pixel in Meter, ist direkt möglich. Eine Bildauswertung kann somit unabhängig von einem Sensormodell und photogrammetrischen Aspekten erfolgen.

Neben den Multispektraldaten kommen *Digitale OberflächenModelle*, DOMs, mit einer typischen Auflösung von ein bis zwei Metern zum Einsatz. Sie werden photogrammetrisch aus den HRSC-Daten abgeleitet. Ein DOM ist ein Bildraster, das jedem Bildpunkt eine Oberflächenhöhe über einem vorgegebenen Ellipsoid zuordnet. Die Oberfläche erfasst also nur die jeweils höchsten Punkte. So wird bei einem Gebäude die Höhe des überstehenden Dachs, eines Balkons oder Erkers erfasst, auch wenn dieser Gebäudeteil nicht bis zum Boden hin fortgesetzt ist. Aus diesem Grund nennt man ein DOM ein 2,5-D-Modell. Im Gegensatz zu einer 3-D-Repräsentation, bei der jeder Punkt im dreidimensionalen Raum attribuiert werden kann, ist dies bei einem DOM nur für die Oberfläche aller Objekte möglich.

Ein DOM stellt damit immer die Vereinigung von Grund- und Objekthöhe dar. Die Bestimmung von Objekthöhen ist so nur bei ebenem Grund eine triviale Aufgabe. Sobald auch das unterliegende Gelände Höhenvariationen aufweist, gilt es, diese von objektbedingten Höhenvariationen zu separieren. Besonders durch die zunehmende Verbreitung von Höhendaten aus Laser-Scannern ist diese Thematik in den Mittelpunkt zahlreicher Untersuchungen geraten. Vielfach wird die Bestimmung der Objekthöhe bezüglich der Grundhöhe als Normieren des Oberflächenmodells bezeichnet. In Kapitel 3 werden zwei Ansätze hierzu entwickelt.

Nach diesem Normierungsschritt sind Oberflächenmodelle prinzipiell hervorragend zur Extraktion erhöhter Objekte geeignet. Bei den stereophotogrammetrisch gewonnenen HRSC-DOMs stellt die genaue Erfassung von Höhengsprüngen ein Problem dar (s. Abschnitt 2.5.2). Die Bildkanten der Orthobilder an diesen Höhengsprüngen sind jedoch meist von guter Qualität.

## 1.3 Literaturüberblick

Die Objektextraktion stellt das einleitend herausgestellte Bindeglied zwischen Bilddaten und interpretierter Geoinformation dar. Der weit gefasste Begriff der „Objektextraktion“ wird in dieser Arbeit verwendet, um zahlreichen damit verbundenen Themenbereichen gerecht zu werden. Konkret werden im Rahmen dieser Arbeit folgende Themen der Bildverarbeitung berührt: Bestimmung/Interpolation der Grundhöhe aus dem DOM, die DOM-Normierung, die Extraktion erhöhter Objekte, die Konturextraktion, Raster-zu-Vektor-Konvertierung, Vegetationssegmentierung, Baumdetektion, Vegetationsindizes, multispektrale Klassifikation, Schattendetektion, Gebäude- bzw. Hausdetektion, Kantendetektion, Konturadaption bzw. Aktive Konturen. Statt der genannten Begriffe sind selbstverständlich auch andere sinnverwandte Bezeichnungen üblich.

Dieser Abschnitt dient der überblicksartigen Einordnung des nachfolgend in Abschnitt 1.4 präsentierten Vorgehens. Da zu den in dieser Arbeit vorgestellten Teilschritten meist mehrere verwandte Ansätze existieren, wäre eine Betrachtung dieser Methoden zu detailliert. Darauf wird in den betreffenden Kapiteln in Form eines Literaturüberblicks zur jeweiligen Thematik eingegangen.

BRUNN (1998) und MAYER (1999) geben einen Überblick über Ansätze zur automatischen Gebäudeextraktion aus Luftbildern<sup>1</sup>. Die Herangehensweisen werden nach unterschiedlichen Kriterien wie z. B. Datenkomplexität oder dem verwendeten Gebäudemodell gruppiert. Auch nachfolgend werden in Abschnitt 1.3.1 unterschiedliche Aspekte einer Objektextraktion schrittweise diskutiert.

Im Anschluss daran werden in Abschnitt 1.3.2 Ansätze zur Gebäudeextraktion, die dieser Arbeit methodisch ähnlich sind, vorgestellt und bewertet.

### 1.3.1 Grundlegende Konzepte der Gebäudeextraktion

Verfahren zur Extraktion künstlich errichteter Objekte (z. B. Gebäude oder Straßen) orientieren sich meist an geometrischen Merkmalen wie Geradlinigkeit oder rechten Winkeln. Durch die mit der Bildaufnahme einhergehenden Projektion auf eine zweidimensionale Bildebene wie auch durch die nachfolgenden Entzerrungsschritte wird die Geometrie der abgebildeten Objekte beeinflusst. Erst durch eine geeignete *geometrische Modellierung des Sensors* ist es möglich, Aussagen über die geometrischen Beziehungen der Bildobjekte zu machen. Zur Geokodierung der Bildinhalte ist dieser Schritt ohnehin unerlässlich. Eine Sensormodellierung anhand vorliegender Bilddaten kann aller-

---

<sup>1</sup>Der Begriff „Luftbild“ wird in dieser Arbeit synonym für alle bildgebenden Verfahren der luftgestützten Erdfernerkundung verwendet, also nicht nur für die klassische Fotografie.

dings nur als Notbehelf verstanden werden. Sie wird nie die Genauigkeit von Labormessungen und einer Positions- und Lageerfassung während des Bildflugs erreichen.

Die Arbeiten von MCGLONE UND SHUFELT (1994) sowie SHUFELT (1996) basieren beispielsweise auf schrägblickend erfassten Bilddaten. Die in den Bildern detektierten Linien können nur durch eine explizite Sensormodellierung sinnvoll zueinander in Beziehung gesetzt werden.

Die zwei letztgenannten Ansätze gehören auch zur Gruppe der Verfahren, denen eine explizite *Liniendetektion* vorgeschaltet ist. Durch die Liniendetektion wird der Großteil der linearen Bildkanten erfasst und liegt dann in Form kleiner Liniensegmente vor. Ziel des Auswerteprozesses ist die geeignete Kombination der Liniensegmente sowie ggf. deren Verlängerung oder Verkürzung. Die Zuordnung der Liniensegmente zu einem Objekturnriss muss wegen der riesigen Anzahl der kombinatorisch möglichen Hypothesen anhand weiterer Faktoren stark eingegrenzt werden. Dazu werden geometrische Kriterien (s. z. B. SHUFELT 1997), die nachfolgend noch beschriebenen Gebäudemodelle (s. z. B. NEVATIA ET AL. 1997, LIN UND NEVATIA 1998), DOM- bzw. Stereoinformation (s. z. B. KIM ET AL. 2000) oder bereits vorliegende ungefähre Objektkonturen (YOON ET AL. 1999) verwendet. Bei diesem in zwei Stufen aufgeteilten Vorgehen (Liniendetektion, Kombination) *summieren sich Ungenauigkeiten und Fehler auf*: Beispielsweise tauchen Ungenauigkeiten in der Liniendetektion nicht nur im Endresultat auf, sondern können auch zu falschen Entscheidungen bei den nachfolgenden Schritten führen. Da der Liniendetektor in den meisten Fällen *nur lineare Strukturen* erkennt, ist auch die darauf aufbauende Objektextraktion im Hinblick auf die Objektgeometrie beschränkt. Nicht-lineare Bildkanten werden nicht oder nur bruchstückhaft erfasst und können in der Zuordnungsphase auch nicht mehr rekonstruiert werden.

Als Vertreter eines linienbasierten Vorgehens seien KRISHNAMACHARI UND CHELLAPA (1996), SHUFELT (1997), YOON ET AL. (1999), SOHN UND DOWMAN (2001) genannt. Neben Linien detektieren MCKEOWN (1990) zusätzlich Eckpunkte zur Gebäudeextraktion.

In mehreren Ansätzen ist der eigentlichen Auswertung die *Konstruktion eines Oberflächenmodells* aus Multiblickwinkeldaten vorgeschaltet (s. z. B. HERMAN UND KANADE 1984, ROUX UND MCKEOWN 1994, COLLINS ET AL. 1995, FÉRE ET AL. 1997). Eine Höhenbestimmung durch Zuordnen von Objektprimitiva erfolgt z. B. bei FISCHER ET AL. (1998). Die Gruppierung dieser 3-D-Primitiva führt dann zu Hypothesen z. B. für Gebäudedächer (KIM ET AL. 2000).

Mit *modellbasierten Ansätzen* wird objektspezifisches Wissen in die Gebäudeextraktion einbezogen. Die Form eines Gebäudes soll dabei durch mehr oder weniger komplexe parametrisierte Modelle erfasst werden; die jeweiligen Größen werden aus den Bilddaten oder auch aus Höhendaten abgeleitet. So werden zum Beispiel anhand der mittleren quadratischen Abweichung des auf die Bildebene projizierten Modells von den Bilddaten passende Parameter gesucht, das Modell also in die Bilddaten eingepasst. Meist kommt eine ganze Bibliothek von Modellen zum Einsatz, aus der dann das am besten passende für das konkrete Objekt gewählt wird. Die Komplexität der in der Literatur verwendeten Modelle ist je nach Ansatz unterschiedlich. So arbeiten VINSON ET AL. (2001) mit Rechtecken, die mittels Aktiver Konturen („Deformable Templates“) an den Gradienten des zugehörigen Orthobilds

angepasst werden. LIN UND NEVATIA (1998) verwenden als zweidimensionales Gebäudemodell Parallelogramme, die dann zu komplexeren Objekten zusammengesetzt werden. Planare Flächen im dreidimensionalen Raum kommen beispielsweise bei BIGNONE ET AL. (1996) und KIM UND MULLER (1998) zum Einsatz. Um Gebäude mit Flachdach und rechteckiger Grundfläche zu modellieren, verwenden COLLINS ET AL. (1998) und SCHUTTE (1996) Quader. Viele Haustypen können bereits mit Prismen modelliert werden, wie durch WEIDNER (1997) beschrieben. Gebäudemodelle mit Dachneigung, je nach Implementierung mit einer unterschiedlichen Anzahl an Dachfirsten, z. B. das Satteldach oder das Walmdach, werden bei GRÜN UND DAN (1997), HELLER UND QUAM (1997), SPREEUWERS ET AL. (1997), LÄBE UND GÜLCH (1998) und GÜLCH ET AL. (1999) eingesetzt. JAYNES ET AL. (1997a,b) verwenden Oberflächenprimitive wie Ebenen, Zylinder, Satteldach oder Gabeldach, um die entsprechenden Oberflächennormalen mit denen aus einem Höhenmodell abzugleichen.

Allgemein lässt sich sagen, dass mit komplexeren Modellen auch eine größere Vielfalt an Gebäuden erfasst werden kann. Jedoch werden für eine verlässliche Schätzung der Modellparameter auch mehr Evidenzen aus dem Bild bzw. Oberflächenmodell benötigt. Je nach Auflösung der Bilddaten macht die Modellierung bestimmter Details somit keinen Sinn mehr. Die Rekonstruktion von Gebäuden anhand sehr allgemeiner Modelle kann allerdings durch die Integration von *externem Wissen* stabilisiert werden. So verwenden HAALA UND BRENNER (1997, 1998) allgemeine Polyeder als Gebäudemodelle; zusätzlich wird dabei auf einen (zweidimensionalen) Gebäudegrundplan zurückgegriffen.

Da sich die vorliegende Arbeit von linien- und modellbasierten Ansätzen abgrenzt, werden nachfolgend die charakteristischen Eigenschaften dieser Ansätze zusammengetragen.

Wird ein DOM verwendet, so sind bei beiden Vorgehensweisen Implementierungen *ohne vorhergehende Normierung des DOM* möglich. Das Einpassen eines Modells kann auf anderen Merkmalen als den absoluten Höhenangaben erfolgen, wie den Oberflächennormalen oder Höhenkanten. Linienbasierte Ansätze gehen per se von Kanten aus. Voraussetzung für die Verlässlichkeit dieser Merkmale ist jedoch ein *hochgenaues DOM*. Wie in Abschnitt 2.5.2 gezeigt wird, ist das stereophotogrammetrisch gewonnene HRSC-DOM jedoch oft gerade an Höhengsprüngen mit Ungenauigkeiten behaftet.

Sowohl bei linien- wie auch bei modellbasierten<sup>2</sup> Ansätzen ist eine implizite *Trennung von Gebäuden und Bäumen* möglich. Bäume weisen keine geradlinigen Kanten und keine gleichmäßig verteilten Oberflächennormalen auf und scheiden daher bei der Gebäudeextraktion aus.

Beide Konzepte weisen *Einschränkungen bezüglich der Gebäudeform* auf. Deren Ursachen und Auswirkungen wurden in diesem Abschnitt bereits diskutiert.

Linienbasierte Ansätze haben durch die Trennung von Liniendetektion und -zuordnung das ebenfalls schon herausgestellte Problem der *Fehlerkumulation*.

Beide Herangehensweisen wurden speziell zur Gebäudeextraktion entwickelt. Eine Erweiterung

---

<sup>2</sup>Bei der Verwendung einfacher Gebäudemodelle.

auf die *Extraktion anderer kartografisch relevanter Objekte*, wie Bäume oder Straßen, ist *nur schwer möglich*.

Abschließend seien *vereinfachende Annahmen* genannt, die bei unterschiedlichen Ansätzen aufgabenspezifisch oder zu Testzwecken vorausgesetzt werden. Bei ebenem Gelände kann auch bei der Verwendung absoluter Höhenmaße von einer Normierung des DOM auf Grundhöhe abgesehen werden. Die erhöhten Regionen sind in diesem Fall schon durch Anwendung eines Schwellwerts zu bestimmen (s. z. B. EHLERS ET AL. 2002).

Da die Untersuchungsgebiete von MASON UND BALTSAVIAS (1997) Hüttensiedlungen mit wenig Vegetation sind, wird dort auf eine explizite Trennung erhöhter Regionen in Bäume und Gebäude verzichtet.

Die meisten Ansätze zur Gebäudeextraktion funktionieren zufriedenstellend bei freistehenden Häusern. So bestehen die Testdatensätze einiger Publikationen aus Aufnahmen kleinerer Einzelhaussiedlungen (beispielsweise der letztgenannte Ansatz von MASON UND BALTSAVIAS (1997) zur Hüttendetektion aus dem internationalen Ascona-Testdatensatz der ETH Zürich, s. GRÜN ET AL. (1995)). Sind mehrere Häuser in Form eines Gebäudeblocks direkt benachbart, stellt die separate Behandlung der Häuser ein Problem dar. Denn gerade bei den in Großstädten typischen Gebäudeblocks haben die Einzelhäuser oft eine ähnliche Dachform und bestehen aus ähnlichen Dachmaterialien. Ansätze, die sich an diesen Kriterien orientieren, sind somit eigentlich nur zur Extraktion von Hausgruppen geeignet.

### 1.3.2 Modellfreie Ansätze zur Gebäudeextraktion

Im letzten Abschnitt wurden Konzepte der Gebäudeextraktion im Hinblick auf eine allgemeine Abgrenzung vorgestellt. Aufgrund der herausgestellten Problematik modellbasierter Ansätze wird nachfolgend genauer auf drei modellfreie Herangehensweisen zur Objektextraktion eingegangen, die in mehreren Punkten dieser Arbeit ähnlich sind. Unter *Modellfreiheit* wird in dieser Arbeit das Fehlen eines explizit definierten Gebäudemodells verstanden.

VESTRI UND DEVERNAY (2001) nutzen für ihre Erfassung der Gebäudekonturen ausschließlich ein DOM. Wie in der vorliegenden Arbeit werden zuerst erhöhte Regionen aus einem auf Grundhöhe normierten DOM extrahiert. Die erhöhten Regionen werden dann durch das Einpassen planarer Flächen approximiert und geglättet. Aus dem so segmentierten DOM wird eine polygonale Gebäudebeschreibung abgeleitet. Dabei wird bereits die Position der Polygonstützpunkte in Bezug auf die Nachbarn und den Verlauf der Konturlinie optimiert. Anschließend findet eine Feinkorrektur der Polygonmodelle mittels geometrischer Nebenbedingungen, die die Rechtwinkligkeit oder Geradlinigkeit fördern sollen, statt.

SOHN UND DOWMAN (2001) stellen einen kanten- und regionenbasierten Ansatz zur Gebäudedetektion lediglich anhand von monochromatischen Bilddaten im Ein-Meter-Bereich vor. Sie nutzen vorab detektierte Linien, allerdings nicht direkt. Aus der Richtungsinformation der Linien werden

die Hauptorientierungen bestimmt. Dann folgt eine Übersegmentierung des Bildes in Segmente, die durch Polygone mit Kantenausrichtungen ausschließlich gemäß den Hauptorientierungen begrenzt sind. Die Kanten dieser Polygone stammen so alle von Bildkanten oder deren geradliniger Verlängerung ab. Die Polygone werden anschließend aufgrund eines Homogenitätskriteriums in einem Regionenwachstumsprozess zu Gebäudesegmenten verschmolzen. Eine Übereinstimmung der Gebäudesegmentkonturen mit Bildkanten sowie deren Geradlinigkeit ist somit implizit gegeben.

Die beiden genannten Autorengruppen extrahieren Gebäudekonturen durch die Betrachtung geometrischer Formmerkmale. Eine Erweiterung der Gebäudeextraktion auf eine Extraktion mehrerer kartografisch relevanter Objektkategorien ist so nur schwer möglich. Ein hierarchisch strukturiertes Vorgehen, bei dem schrittweise unterschiedliche Objektattribute überprüft werden, bietet sich für eine solch verallgemeinerte Objektextraktion an. Durch die aufeinanderfolgende Anwendung kann der Objektbegriff immer weiter verfeinert werden. In einem baumartig strukturierten Vorgehen zur Objektextraktion entspräche die Methodik zur Gebäudeextraktion einem der Äste. Die Extraktion anderer Objekte wird durch weitere Äste repräsentiert. Viele der betrachteten Kriterien können so mehrfach genutzt werden.

STRAUB UND HEIPKE (2001), GERKE ET AL. (2001) stellen ein hierarchisch strukturiertes Vorgehen zur Bildanalyse vor, in dem der Objektbegriff in fünf Abstraktionsstufen verfeinert wird. Die oberste Ebene bildet der Begriff „Landschaft“. Dieser untergliedert sich in „Superklassen“ wie Siedlung, offene Landschaft, Wald und Wasser. Als nächste Abstraktionsebenen folgen „Gebiete“ wie Baumreihe, Baumgruppe, Gebäuderegion oder Straßennetz. Auf der „Objektebene“ sollen einzelne Bäume, Häuser oder Straßensegmente erkannt werden. Schließlich folgt eine weitere, feiner untergliederte „Komponentenebene“. Sie umfasst Baumkronen, Dächer oder Straßenmarkierungen. Ausgangsdaten sind Colorinfrarotbilder (CIR) und normierte Oberflächenmodelle (nDOMs). Sowohl zur Gebäude- als auch zur Baumextraktion werden aus dem nDOM erhöhte Regionen bestimmt. Aus den CIR-Bildern werden Vegetationssegmente über einen Vegetationsindex (NDVI) ermittelt. Die Verschneidung der beiden Resultate führt zu Baumgruppen und Gebäuderegionen. Die durch das geringer aufgelöste nDOM bedingte grobe Erfassung der Gebäuderegionen wird anhand des NDVI-Kriteriums überarbeitet. Ausgehend von den Zentren der Gebäuderegionen wird die Segmentform durch Vorgabe eines NDVI-Wertebereichs neu ermittelt. In die daraus resultierenden Regionen werden momentbasiert Rechtecke eingepasst. Die Modellierung mit Rechtecken wird immer feiner gestaltet, so dass auch komplexe Gebäudeformen erfasst werden können. Im Sinne dieser Arbeit ist der Ansatz trotz des rechteckbasierten Vorgehens somit als modellfrei zu betrachten.

Das momentbasierte Einpassen der Rechtecke ist problematisch, wenn die Segmentform zwei ähnlich starke Hauptachsen besitzt, also bei Quadraten oder Kreisen, oder zu komplex ist. Der Einpassvorgang basiert nur auf den überarbeiteten Gebäuderegionen, ist also nur indirekt von der eigentlichen Bildinformation und Bildkanten abhängig. Lediglich die vorhergehende NDVI-basierte Segmentüberarbeitung kann Bildkanten einbeziehen. Dabei wird von einer möglichst homogenen NDVI-Verteilung ausgegangen. Doch die in hochauflösenden Bilddaten sichtbaren Dachstrukturen und un-

terschiedliche Dachmaterialien beeinflussen den NDVI und können so zu einer *unvollständigen oder fehlerhaften Erfassung des Gebäudesegments* führen. Ferner ist durch die Wahl einer Kernregion pro Segment für die Segmentüberarbeitung *keine Unterscheidung mehrerer Häuser eines Gebäudekomplexes* möglich. Entweder wird durch die Segmentüberarbeitung der gesamte Gebäudekomplex oder aber nur *ein* Teil davon erfasst.

### 1.4 Ziel und Aufbau der Arbeit

Im Überblick über bestehende Ansätze zur Gebäudeextraktion des vorhergehenden Abschnitts wurden mehrere Unzulänglichkeiten der Verfahren herausgestellt. Eine Beseitigung dieser Kritikpunkte wird in der nun folgenden Zielstellung für diese Arbeit gefordert. Anschließend wird das Vorgehen zum Erreichen dieser Ziele, ein hierarchisch strukturierter Segmentierungsansatz, präsentiert. Der Abschnitt endet mit einer Vorstellung der Struktur der Arbeit, in der die Inhalte der einzelnen Kapitel skizziert werden.

#### 1.4.1 Zielsetzung

Wie einleitend in Abschnitt 1.1 herausgestellt, ist die Automatisierung der Bildinterpretation im Hinblick auf den Markt der Geoinformationssysteme zentrale Motivation für diese Arbeit. Daraus ergeben sich folgende Anforderungen: Das Vorgehen zur Objektextraktion muss *automatisiert* von Bilddaten zu *geometrisch genauen Objektinformationen* wie Konturlinien oder zugeordneten Attributen führen. Die Resultate müssen in einem gängigen *GIS-Datenformat* vorliegen. Um einen möglichst breiten Einsatzbereich abzudecken, sollte das Vorgehen mit *möglichst wenig Vorgaben* oder Annahmen bezüglich der Objekte oder der Bildinhalte auskommen. Ebenso sollte die *Erweiterbarkeit* der Objektextraktion von Gebäuden zu allgemeineren Objektkategorien und die *Nutzbarkeit von Teilschritten* für unterschiedliche Aufgabenstellungen angestrebt werden.

Datengrundlage der Arbeit sind nach Abschnitt 1.2 hochauflösende multispektrale Orthobilder sowie ein Oberflächenmodell. Weitere Informationsquellen, wie bereits vorhandene Segmentierungsergebnisse, Bebauungspläne oder topografische Karten, können die Extraktionsergebnisse verbessern. Da von deren Vorliegen jedoch nicht ausgegangen werden kann, soll die zu entwickelnde Methodik *unabhängig von externer Information* funktionieren.

Die *hohe geometrische Auflösung* der Ausgangsdaten und die *komplementäre Information* von Multispektraldaten und Oberflächenmodell sollen im Extraktionsvorgang genutzt werden. Die Methodik sollte prinzipiell auf andere Datensätze, die aus Multispektral- und Oberflächendaten bestehen, *übertragbar* sein. HRSC-spezifische Besonderheiten sollten herausgestellt werden. Bei der Implementierung der Methoden ist die mit hochauflösenden Bilddaten verbundene hohe Datenmenge zu berücksichtigen.

Neben der für ein Stadt-GIS bedeutenden Objektkategorie der Gebäude gibt es zahlreiche weitere



GIS-relevante Objekte, die auf Luftbildern erkennbar sind. Die eingangs geforderte Erweiterbarkeit des Extraktionsansatzes erhöht den Anwenderkreis und ermöglicht eine effiziente Nutzung der Bildinformation bzw. der daraus abgeleiteten Zwischenergebnisse.

Die Nutzbarkeit von Teilschritten ist ebenfalls durch ein erweitertes Anwendungsspektrum motiviert. Auf weitere Vorzüge eines modularen Aufbaus wird im folgenden Abschnitt hingewiesen.

Je weniger Vorgaben an die zu extrahierenden Objekte oder die Bildszene gestellt werden, desto höher ist die Nutzbarkeit der Gebäudeextraktion. Um eine breite Palette von Gebäudetypen bearbeiten zu können, müssen Festlegungen bezüglich der Form der Gebäude, wie sie bei modellbasierten Ansätzen üblich sind, vermieden werden. Auch eine Konzentration auf bestimmte Dachmaterialien wie Dachpappe oder rote Ziegeln ist zu einschränkend.

Die Extraktion sollte ferner möglichst unabhängig von der Morphologie des zugrunde liegenden Geländes sein. Da das Interesse an automatisiert gewonnener Geoinformation im Stadtbereich besonders hoch ist, muss auch auf die Anwendbarkeit in dicht besiedelten Regionen geachtet werden. Weiterhin sollten direkt aneinandergrenzende Häuser bzw. Häuser eines Häuserblocks separat extrahiert werden.

### 1.4.2 Vorgehen zur Objektextraktion

In diesem Abschnitt wird ein Überblick über den Ablauf der in den nachfolgenden Kapiteln ausgeführten Schritte zur Objektextraktion gegeben. Zunächst ist dazu der Objektbegriff zu klären. Welche Objekte sind von Interesse? Welche Attribute eines Objekts sollen durch die Extraktion bestimmt werden?

Der Schwerpunkt dieser Arbeit liegt auf der Extraktion von Häusern. Dazu werden folgende gebäudetypische Attribute benutzt: Gebäude weisen eine *Mindesthöhe* auf und sie unterscheiden sich in der *spektralen Zusammensetzung* des reflektierten Lichts von Vegetation. Zudem sind selbst direkt benachbarte Häuser visuell als *separate Einheiten* auszumachen. Neben unterschiedlichen Dachmaterialien ist der Übergang aneinandergrenzender Häuser durch Brandmauern gut zu erkennen. Schließlich wird die Form der Haussegmente betrachtet. Zu *gebäudetypischen Geometrien* zählen geradlinige und zueinander rechtwinklige Konturlinien.

Für das in dieser Arbeit vorgeschlagene Vorgehen werden diese Attribute schrittweise überprüft und zur Selektion der vorhandenen Segmente benutzt. In einer *hierarchisch geordneten Abfolge* der Extraktionsschritte wird der Objektbegriff in jeder Verarbeitungsstufe feiner aufgelöst: Aus der anfänglich vorhandenen Menge aller Bildpunkte werden *erhöhte Regionen* selektiert. Diese Regionen werden aufgeteilt in *Baumregionen und Gebäude*. Nachfolgend werden die Gebäuderegionen unterteilt in kleinere Einheiten, möglichst *einzelne Häuser*. Einzelne Häuser stellen die unterste Ebene der Objekthierarchie dar.

In dieser Arbeit wird von großen Segmenten (ursprünglich das gesamte Bild) ausgegangen, die immer feiner unterteilt werden. Die Segmentierungsschritte von SOHN UND DOWMAN (2001) (s. Ab-

schnitt 1.3.2), in denen aus kleinen Segmenten ein größeres Objektsegment zusammengefügt wird, verlaufen invers.

Das ebenfalls in Abschnitt 1.3.2 skizzierte Vorgehen von STRAUB UND HEIPKE (2001), GERKE ET AL. (2001) basiert auch auf mehreren Abstraktionsstufen. Die unterste Abstraktionsebene, die „Komponentenebene“ (Baumkronen, Dächer oder Straßenmarkierungen), wäre von der Bedeutung her identisch mit der Objektebene dieser Arbeit, da sich aus der Orthoperspektive Bäume nur durch ihre Baumkrone, Häuser nur durch ihre Dächer ausmachen lassen<sup>3</sup>.

Die Aufgliederung der Segmente im Lauf der hierarchischen Objektextraktion hat das Zustandekommen anderer sinnvoller Klassen als der im Vordergrund stehenden Gebäude zur Folge. So ergeben sich aus der Negation der jeweils betrachteten Objekte niedrigere Regionen oder Baumregionen<sup>4</sup>. Eine Extraktion anderer Objekte wie z. B. Rasenflächen oder einzelner Bäume kann so durch eine hierarchisches Vorgehen analog zur Gebäudeextraktion erfolgen. Je nach Anwendung ist das Prinzip der hierarchischen Verfeinerung also *auf beliebige Objektklassen zu übertragen*. Für eine Differenzierung zahlreicher Objektkategorien bietet sich der Aufbau eines hierarchischen Extraktionssystems geradezu an. Hierbei kommt komplementäre Bildinformation voll zum Tragen. Auch die jeweils anfallenden Zwischenergebnisse werden durch die mehrfache Verwendung optimal genutzt. Eine an die Extraktionsphase anschließende Klassifikation innerhalb einer Objektkategorie kann dann zwischen unterschiedlichen Ausprägungen einer Objektkategorie unterscheiden. Als Beispiel hierfür sei die Differenzierung einzelner Bäume in verschiedene Baumklassen genannt. Das in Abschnitt 1.1 erwähnte objektbasierte Vorgehen (S. 5) bietet sich hierfür an.

Nach einer Klassifikation der Objektsegmente ist der Übergang von Pixelinformation zu abstrakten Objektbeschreibungen vollzogen. Statt der nicht-interpretierten Bildinformation können in einem geografischen Datenbanksystem so kompakte Objektbeschreibungen wie Position, Umriss, Kategorie und Klasse oder Ausprägung sowie die im Lauf der Extraktionsphase ermittelten Attribute abgelegt werden.

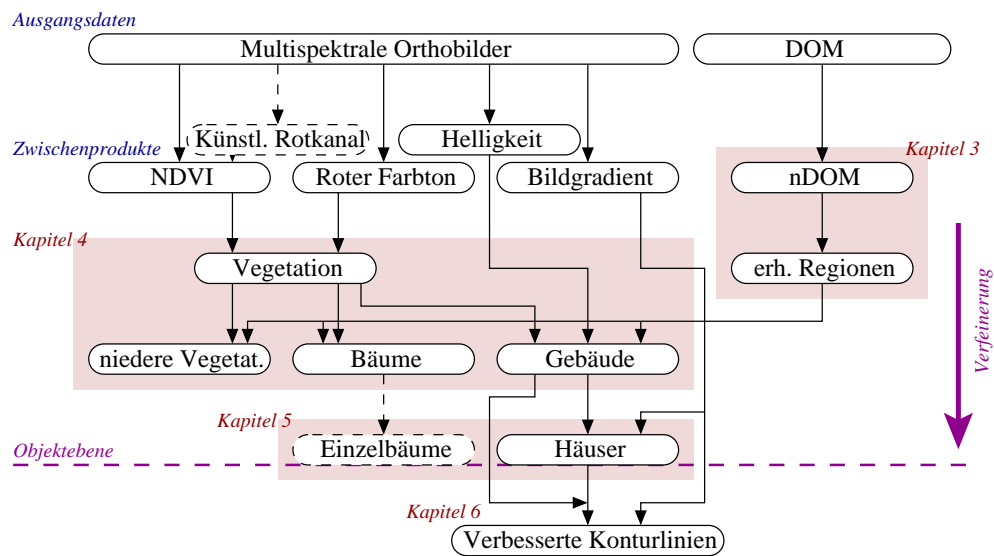
Das in dieser Arbeit vorgeschlagene Vorgehen zur Objektextraktion ist in Abbildung 1.3 schematisch wiedergegeben. Aus den Ausgangsdaten werden Zwischenprodukte abgeleitet. Zum einen gibt es Zwischenprodukte, die vor Beginn aller Extraktionsschritte ermittelt werden können, wie beispielsweise die Helligkeit oder der Bildgradient. Zum anderen erfordern die Module zur Objektextraktion Vorverarbeitungsschritte, die abhängen von den vorab erzielten Ergebnissen. Diese Schritte sind in Abbildung 1.3 nicht erfasst.

Durch die Normierung des Oberflächenmodells auf Grundhöhe nach Kapitel 3 ist es möglich, *unabhängig von der Geländemorphologie* die erhöhten Regionen zu extrahieren. Anhand der multispektralen Information wird in Kapitel 4 Vegetation detektiert. Prinzipiell führt eine Vereinigung mit der Segmentierung erhöhter Regionen dann zu der Differenzierung in Bäume, Gebäude, niedrigere Ve-

---

<sup>3</sup>Von der Möglichkeit, aufgrund der Schatteninformation auf die Form eines Objekts zu schließen, sei hier abgesehen.

<sup>4</sup>Wie nachfolgend noch ausgeführt wird, resultieren die Gebäudesegmente bei der in Kapitel 4 behandelten Unterteilung der erhöhten Regionen tatsächlich aus dem Ausschluss der vorab bestimmten Baumsegmente.



**Abbildung 1.3:** Schematischer Überblick des hierarchischen Vorgehens zur Objektextraktion mittels multispektraler Orthobilder und eines DOM.

getation und sonstige niedere Regionen. Durch dieses Vorgehen unterliegen die Gebäudesegmente *keiner Beschränkung bezüglich ihrer Form oder Geometrie* und nur einer *minimalen Einschränkung bezüglich der spektralen Charakteristika*. Die nachfolgende Aufteilung der Gebäuderegionen in einzelne Häuser gemäß Kapitel 5 orientiert sich an den durch Brandmauern hervorgerufenen Helligkeitswechseln. Die letztgenannte Freiheit bezüglich Form und Material der Häuser bleibt also erhalten. Mit dem Vorliegen einzelner Häuser ist dann die unterste Ebene der hierarchischen Extraktion erreicht. Da die Haus- oder Gebäudesegmente bereits als solche erkannt sind, können nun gebäudespezifische Konturlinien forciert werden. Dieser Überarbeitung der Konturlinien widmet sich Kapitel 6. Im nachfolgenden Abschnitt wird genauer auf die Inhalte der einzelnen Kapitel wie auch ihre Strukturierung im Gesamtkontext eingegangen.

Die Überarbeitung der Konturlinie von Häusern bildet den Abschluss des in dieser Arbeit vorgestellten Ansatzes zur Gebäudeextraktion. Resultat ist eine Konturbeschreibung im zweidimensionalen Raum. Die vorhergehenden Schritte gewährleisten, dass die Gebäudesegmente eine Mindesthöhe aufweisen. Eine Attribuierung der Gebäudefläche mit Höheninformation aus dem DOM bzw. nDOM führt zu einer 2,5-D-Beschreibung der Gebäude.

Die Ausgangsdaten sind Standarddatenprodukte der HRSC-Datenverarbeitung und werden geokodiert als ein zusammengehöriger Datensatz angeboten. Eine aufwändige Berechnung des Oberflächenmodells durch Zuordnung von Merkmalen in Bilddaten mehrerer Blickwinkel kann so entfallen. Weiterhin ist durch das geokodierte Orthoformat zur Auswertung keine explizite Modellierung der Sensorgeometrie nötig. Die hohe geometrische Qualität der Bilddaten (s. dazu Abschnitt 2.4) ermöglicht eine *geometrisch genaue Objektextraktion*.

Durch die Kombination von Information aus multispektralen Daten und dem DOM werden die *jeweiligen Schwachstellen für die Objektextraktion wechselseitig kompensiert* (s. Vorstellung der Ausgangsdaten in Abschnitt 1.2 und Diskussion der Eignung für eine automatisierte Interpretation in Abschnitt 2.5). Das DOM wird zur Extraktion erhöhter Objekte genutzt, die Multispektraldaten zur Vegetationssegmentierung, die gute Kantenqualität der Orthobilder zur Korrektur der Gebäudeumrisse.

Die Methoden zur Objektextraktion, die in den Kapiteln 3–6 diskutiert werden, sind gegliedert in aufgabenspezifische *Module*. Führt man alle Programmmodule hintereinander aus, so gelangt man von den Ausgangsdaten (multispektrale Orthobilder und ein DOM) zu Konturlinien einzelner Häuser. Diese einzelnen Schritte nicht in einen Gesamtablauf zu integrieren (womit eine automatische Objektextraktion gegeben wäre), bietet gerade bei der Umsetzung wissenschaftlicher Erkenntnisse mehrere Vorteile:

- Die Module sind in sich geschlossene Einheiten mit modulspezifischen Vorverarbeitungsschritten. Die Schnittstellen zu dem Vorgänger- und Nachfolgermodul sind über die jeweils berechneten Resultate (Zwischenergebnisse im Gesamtablauf) gegeben. Die Ersetzung eines Moduls durch eine andere Herangehensweise ist so problemlos möglich. In Kapitel 3 werden beispielsweise zwei Methoden zur Berechnung eines Geländemodells aus einem Oberflächenmodell entwickelt. Diese können zu Testzwecken frei ausgetauscht werden. Sollen mehrere Ansätze für dieselbe Problematik miteinander verglichen werden, ist ein modularer Aufbau unabdingbar.
- Die Resultate eines jeden Moduls sind bedeutungstragende Segmentierungsergebnisse. Sie stehen auch für andere Aufgaben der Bildinterpretation als die hier zentral behandelte Gebäudeextraktion zur Verfügung. So werden die bereits erwähnten Grundmodelle (s. Kapitel 3) auch für hydrologische Modellierungen eingesetzt.
- Das Ablegen der Zwischenergebnisse ermöglicht die Fortsetzung zeitintensiver Arbeitsschritte zu einem späteren Zeitpunkt. Ferner können bereits auf Basis der Zwischenergebnisse die Auswirkung unterschiedlicher Parameterkonstellationen einander gegenüber gestellt werden.
- Die Programme sind eine direkte Umsetzung von Erkenntnissen aus Experimenten zur Objektextraktion. Nicht immer können befriedigende Genauigkeiten erzielt werden. Das Ablegen von Zwischenergebnissen ermöglicht dem Nutzer eine manuelle Feinkorrektur der automatisch ermittelten Resultate. Bei einem Gesamtablauf aller Module ohne manuelle Korrekturschritte pflanzen sich die Fehler eines Moduls in den nachfolgenden Modulen fort; niedrigere Genauigkeiten für den Gesamtablauf sind die Folge. Auf den Nutzen einer manuellen Nachbearbeitung wird beispielsweise in Abschnitt 3.2 (Korrektur der Bildpunkte, die als erhöht gegenüber dem Grund eingestuft sind) oder 5.8.2 (Korrektur der Unterteilung eines Gebäudesegments in einzelne Häuser) explizit hingewiesen.

Der Ablauf der in dieser Arbeit präsentierten Routinen wird als *automatisiert* bezeichnet. Prinzipiell beschränkt sich die Nutzerinteraktion auf die Spezifikation von Dateinamen und *Schwellwerten*. Der Automatisierungsgrad ist also sehr hoch. Ist eine schwellwertbasierte Entscheidung analytisch modellierbar, so lassen sich die Schwellwerte automatisch bestimmen. Beispielsweise wird in Abschnitt 3.3 die lokale Grundhöhe anhand des ersten Maximums im lokalen Höhenhistogramm bestimmt. Andere Entscheidungen basieren auf meist nicht explizit definiertem Alltagswissen: Welche Mindesthöhe müssen Häuser oder Bäume aufweisen, um als solche zu gelten? Welche Grundfläche nehmen Häuser mindestens ein? In solchen Fällen wird dem Nutzer ein plausibler Standardwert vorgegeben, der frei verändert werden kann.

Nach einigen Testläufen haben sich für grundlegende Entscheidungen meist vernünftige Parameterkonstellationen bewährt. Diese werden dann direkt in die Steuerrouninen übernommen und können vom Nutzer zur Laufzeit nicht mehr verändert werden. Dadurch wird der Anwender nur mit einem Mindestmaß an Entscheidungen konfrontiert. Da die Steuerrounne der eigentlichen Auswerteroune jedoch alle Schwellwerte als Prozedurvariablen übergibt, kann die Zuordnung in vom Nutzer beeinflussbare und nicht beeinflussbare Größen leicht verändert werden. In Abhängigkeit von den in der Praxis gemachten Erfahrungen können also weitere Stellgrößen für den Benutzer bereitgestellt oder bestimmte Schwellwerte verborgen werden. Daraus resultiert ein System, das mit den Erfahrungen der Nutzer „lebt“. Diese Herangehensweise bietet sich für wissenschaftliche Fragestellungen an.

### 1.4.3 Struktur der Arbeit

Bevor konkrete methodische Aspekte zur Gebäudeextraktion behandelt werden, wird in Kapitel 2 der Sensor HRSC vorgestellt. Alle in dieser Arbeit verwendeten Bildbeispiele stammen von diesem Sensor. Die vorgestellten Extraktionsansätze sind zwar nicht festgelegt auf Daten dieses Sensors, doch sind bei einer automatisierten Auswertung von HRSC-Daten kameraspezifische Aspekte zu berücksichtigen. Bestimmte Effekte, die während der Bildauswertung festgestellt werden, sind bereits auf das Aufnahmeprinzip oder die HRSC-Datenverarbeitung zurückzuführen. Letztendlich stellen die Unzulänglichkeiten des Oberflächenmodells eine grundlegende Motivation für die vorgestellten Schritte zur Gebäudeextraktion und Konturverbesserung dar.

Die Kapitel 3 bis 6 stellen den eigentlichen Kern der Arbeit dar. Hier wird die vorgeschlagene Methodik zur Gebäudeextraktion behandelt. Diese Kapitel sind entsprechend den schrittweise durchgeführten Teilaufgaben modular aufgebaut. Auf das Thema „Modularität“ ist bereits im vorhergehenden Abschnitt 1.4.2 genauer eingegangen worden. Die Reihenfolge der Kernkapitel entspricht der Abfolge der Schritte zur Extraktion von Gebäuden. Je nach Aufgabenstellung kann die Durchführung der in den Kapiteln 5 und 6 beschriebenen Abläufe auch umgekehrt erfolgen. Die Bedeutung der Abfolge dieser beiden Kapitel wird im Rahmen der Auswertung in Abschnitt 6.7 beleuchtet.

Kapitel 3 bildet die Grundlage für einen Übergang von den Bilddaten zu attribuierten Segmenten. In diesem Kapitel wird ausschließlich auf Basis des Oberflächenmodells gearbeitet. Ziel ist die Ex-

traktion erhöhter Objekte. Dazu werden zwei Methoden zur Normierung eines DOM auf Grundhöhe vorgestellt und miteinander verglichen. Nach Normierung des DOM lassen sich erhöhte Objekte prinzipiell durch Ansetzen einer Höschwelle detektieren.

Diese Regionen erhöhter Objekte werden durch die in Kapitel 4 behandelten Ansätze untergliedert in Bäume und Vegetation. Die Multispektralkanäle der HRSC liefern die dafür benötigte Information. Bedingt durch die andere Auslegung der HRSC-A-Farbkanäle sind im Vergleich zu Daten der HRSC-A-Nachfolgemodelle dafür zusätzliche Schritte notwendig. Nach einer Identifikation der Vegetationsbereiche ergeben sich die Gebäudesegmente prinzipiell aus den verbleibenden erhöhten Regionen.

In Kapitel 5 wird dann ein Vorgehen zur Unterteilung dieser Gebäudesegmente in kleinere Einheiten vorgestellt. Hierbei gilt als Ziel, Segmente von Häuserblocks so aufzuteilen, dass einzelne Häuser resultieren. Als wichtigstes Merkmal wird dabei die durch Brandmauern bedingte Bildkante herangezogen. Aus der Kanteninformation und Ausgangspositionen auf der Mittelachse eines Gebäudesegments werden zahlreiche Kandidaten für eine Unterteilungslinie generiert. Durch schrittweises Einbeziehen unterschiedlicher Plausibilitätskriterien werden falsche Kandidaten effizient aussortiert.

Schließlich wird in Kapitel 6 der Umriss eines Haus- oder Gebäudesegments betrachtet. Die initiale Form des Segments soll so verändert werden, dass dessen Konturlinie möglichst gut mit Bildkanten in der Nähe des Segmentrands übereinstimmt. Dazu wird ein den „Aktiven Konturen“ ähnlicher Ansatz gewählt. Der Prozess zur Anpassung der Konturlinie an Bildkanten wird durch geometrische Nebenbedingungen beschränkt. Dadurch werden gebäudetypische Konturkonstellationen mit rechten Winkeln oder zueinander parallelen Umrisslinien forciert.

Abschließend werden die so gewonnenen Erkenntnisse in Kapitel 7 zusammengetragen. Außerdem wird für die in den Kapiteln 3 bis 6 vorgestellten Methoden ein Ausblick auf Ansätze für weitergehende Forschungsaktivitäten gegeben.

Die Kapitel zur Methodik sind alle nach dem gleichen Prinzip strukturiert: Zuerst wird ein Überblick über bestehende Verfahren zur jeweiligen Problematik gegeben und das eigene Vorgehen motiviert. Nach Vorstellung der jeweils nötigen Vorverarbeitungsschritte folgt die Präsentation der eigentlichen Vorgehensweise. Abschließend werden die Methoden anhand von Testdatensätzen in Bezug auf ihre Laufzeit und die erzielbare Genauigkeit bewertet. Dieser gleichartige modulare Aufbau der methodischen Kapitel soll die Lesbarkeit der Arbeit fördern.

Die Evaluation der vorgestellten Algorithmen geschieht meist nur an einem oder wenigen Testdatensätzen. Dies ist darin begründet, dass ein entsprechendes Referenzergebnis manuell erarbeitet werden muss. Der dafür notwendige Aufwand ist nicht unbeträchtlich. Schließlich ist der hohe manuelle Aufwand zur Ableitung abstrakterer Information *die* zentrale Motivation für diese Arbeit. Durch den eingeschränkten Umfang der Stichproben zur Auswertung sind die resultierenden Angaben nur als Richtwerte zu betrachten. Je nach verwendeten Bilddaten ist mit Schwankungen in Bezug auf diese Angaben zu rechnen.

Zur Visualisierung der Resultate der jeweils behandelten Arbeitsschritte werden die Routinen an

Bildbeispielen angewendet. Um eine gewisse Durchgängigkeit zu erreichen, wird meist auf Bildausschnitte *eines* Datensatzes zurückgegriffen. Je nach Aufgabenstellung bietet sich in Einzelfällen jedoch die Verwendung eines anderen Datensatzes an. So kann die in Kapitel 3 behandelte Normierung der Höhenangaben eines Oberflächenmodells plastisch nur an einem Datensatz mit ausgeprägten Geländeneigungen dargestellt werden.

## 1.5 Implementierungsaspekte

In diesem Abschnitt werden Aspekte beleuchtet, die bei der Umsetzung der hier vorgestellten Methoden in Programmcode eine Rolle spielen. Zum einen wird die zum Einsatz gekommene Programmierungsumgebung vorgestellt (Abschnitt 1.5.1), zum anderen wird allgemein auf Anforderungen eingegangen, die hochauflösende Bilddaten an ein Bildverarbeitungssystem stellen (Abschnitt 1.5.2). Abschließend wird in Abschnitt 1.5.3 auf die Bedeutung der Bildkanten als Objektgrenzen eingegangen.

### 1.5.1 Die Programmierungsumgebung

Nachfolgend wird ein Überblick über die Arbeitsumgebung zur Implementierung und zum Test der in dieser Arbeit besprochenen Routinen gegeben. Dies macht ein Einordnen der in den Auswertungsteilen der Kapitel 3–6 durchgeführten Laufzeitmessungen möglich. Die Angaben zur Arbeitsumgebung sollen nicht als Beschränkung auf die verwendete Software oder den verwendeten Rechner verstanden werden. Neben kürzeren Programmlaufzeiten<sup>5</sup> erlauben modernere Rechensysteme auch die Implementierung von bisher als zu zeitaufwändig eingestuften Arbeitsschritten.

Die vorgestellten Methoden zur Objektextraktion sind alle in der von der Firma RSI entwickelten und vertriebenen Programmiersprache IDL („*Image Description Language*“) implementiert. IDL gilt als matrizenorientierte funktionale Hochsprache für die Bildverarbeitung. Die üblichen Programmkonstrukte wie Schleifen, Abfragen, prozedurale Kapselung, Rekursion etc. sind vorhanden. Ebenso werden Konstrukte für eine grafische Benutzerschnittstelle zur Verfügung gestellt. Weiterhin existieren umfangreiche Funktionsbibliotheken für Fragestellungen in der Mathematik und der Bildbearbeitung. Objektorientierte Ansätze sind seit der Version 5 in das Konzept von IDL aufgenommen. Die Übersetzung der Routinen erfolgt standardmäßig beim ersten Aufruf. IDL wird plattformübergreifend angeboten (es gibt Übersetzer und Entwicklungsumgebungen für mehrere Unix-Systeme, für Windows und für den Mac). Die Anwendung benutzereigener Routinen auf einem anderen System kann direkt, ohne Portierung des Programmcodes, vollzogen werden.

Aufbauend auf IDL wird von RSI das Softwarepaket ENVI für die Analyse von Bilddaten in der Fernerkundung vertrieben. Dieses menübasierte System erleichtert dem Anwender die Analyse und

---

<sup>5</sup>Mit einem im März 2003 aktuellen PC werden die in dieser Arbeit angegebenen Laufzeiten bereits auf einen Bruchteil reduziert.

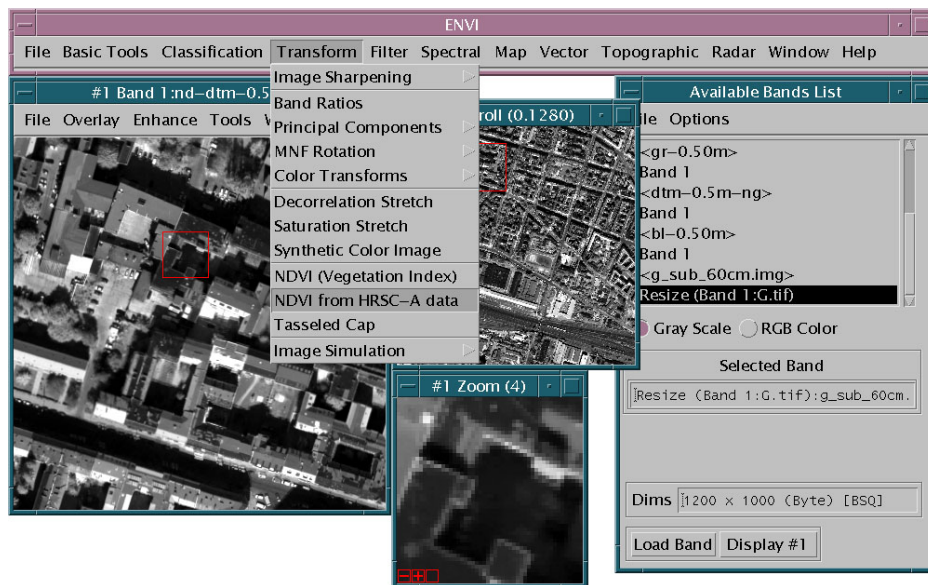


Abbildung 1.4: Die ENVI Arbeitsumgebung.

Bearbeitung von Bilddaten. Dem Programmierer werden weitere, in der Fernerkundung nützliche Routinen zur Verfügung gestellt. Bild- und Vektordaten können in zahlreichen fernerkundungs- und GIS-üblichen Datenformaten gelesen und geschrieben werden. Dies ermöglicht ein Ablegen der Resultate dieser Arbeit in einem GIS-Format.

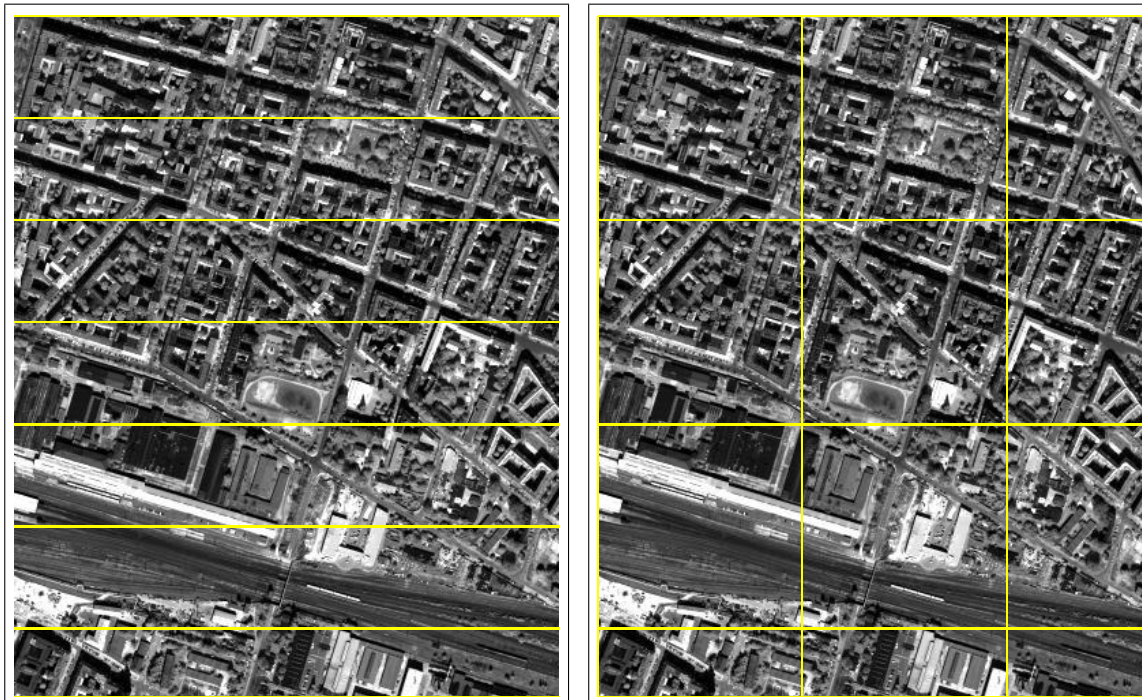
Das ENVI-Menüsystem (s. Abbildung 1.4) lässt sich durch die Integration eigener IDL-Programme flexibel erweitern. Insbesondere zur Nutzung der grafischen Benutzerschnittstelle wird bei den im Rahmen dieser Arbeit implementierten Methoden auf ENVI-Funktionalität zurückgegriffen. Ferner werden ENVI-Funktionen zur Kachelung von Bilddaten verwendet. Auf diesen Aspekt wird in Abschnitt 1.5.2 eingegangen.

Die IDL-Routinen laufen als 32-Bit Anwendung auf dem Betriebssystem Solaris 7 für SPARC-Prozessoren. Als Rechner kommt ein Sun UltraSPARC 10 System mit 256 MB Hauptspeicher zum Einsatz.

### 1.5.2 Anforderungen hochauflösender Daten an die Implementierung von Bildverarbeitungsroutinen

Durch die großflächigen Befliegungsgebiete beanspruchen Bilddaten in der Fernerkundung große Mengen an Speicherplatz. Im Vergleich zu satellitengestützt gewonnenen Aufnahmen umfassen die Bilddaten mit räumlichen Auflösungen unterhalb eines Meters kleinere Zielgebiete; eine typische HRSC-Befliegung deckt die Abmessungen einer Großstadt ab. Durch die hohe Auflösung resultieren jedoch beeindruckende Mengen an Bildinformation. Wie im folgenden Kapitel noch ausführlicher dargestellt wird, weist ein Flugstreifen etwa 5 200 bzw. 12 000 Pixel in der Breite auf. Je nach





(a) Kachelung in einer Dimension

(b) Kachelung in zwei Dimensionen

**Abbildung 1.5:** Bildbeispiel zur Kachelung großer Bilddaten. Einmal wird die Unterteilung nur in einer Dimension durchgeführt (a), das andere Mal in beiden Dimensionen (b).

Aufnahmedauer für den Flugstreifen ergeben sich meist mehrere 10 000 Bildzeilen (im Extremfall mehrere 100 000 Zeilen). Aus mehreren Flugstreifen wird dann ein Gesamtbild mit mehreren 10 000  $\times$  mehreren 10 000 Bildelementen berechnet. Da die HRSC ein Multispektralsensor ist, resultieren vergleichbare Datenmengen für die anderen Bildkanäle.

Diese Abschätzungen sollen veranschaulichen, dass eine Verarbeitung solch großer Bilddaten besondere Vorkehrungen erfordert. Davon, dass man ein Bild vollständig in den Hauptspeicher einlesen kann, um es dann zu bearbeiten, kann nicht ausgegangen werden. Deswegen müssen die Bilder in kleinere Einheiten aufgeteilt werden, die dann sequenziell<sup>6</sup> in den Hauptspeicher eingelesen und bearbeitet werden. Diese Unterteilung wird *Kachelung* („Tiling“) genannt. Im Allgemeinen wird das Bild aufgeteilt in gleich große<sup>7</sup> Rechtecke, wie in Abbildung 1.5(b) veranschaulicht. Die im letzten Abschnitt vorgestellte Programmierungsumgebung ENVI, die zur Implementierung der in dieser Arbeit behandelten Methoden verwendet wird, stellt dazu entsprechende Programmkonstrukte zur Verfügung. Allerdings wird die Bildgröße dabei nur in einer Dimension reduziert, die Bildbreite bleibt erhalten (s. Abbildung 1.5(a)).

<sup>6</sup>Bei Parallelrechnern bietet sich natürlich auch eine parallele Abarbeitung dieser Einheiten an.

<sup>7</sup>Bis auf die Randbereiche des Bildes.

Gerade bei Filteroperationen bietet sich eine Kachelung der zu bearbeitenden Bilder an. Hierbei ist allerdings auf eine ausreichende Überlappung der Kacheln zu achten. Mit einer Überlappung in der Größe des halben Filterkerns wird immer auf real vorhandene Bildinformation zurückgegriffen. Nur so ist ein kontinuierlicher Übergang des Filterungsergebnisses an den Kachelgrenzen gewährleistet.

Im späteren Verlauf des vorgestellten Ablaufs zur Objektextraktion genügt es, ausgewählte Bildbereiche zu behandeln. So erfolgt beispielsweise die Unterteilung der Gebäudesegmente in kleinere Einheiten nach Kapitel 5 separat für jedes Gebäudesegment. Statt durch ein festes Kachelungsraster erfolgt der Bildzugriff durch in der Größe variierende Ausschnittsfenster. Im Fall der Gebäudesegmente entspricht der Ausschnitt dem umgebenden Rechteck des Segments.

Neben der Größe des betrachteten Bildausschnitts ist bei der Entwicklung von Routinen zur Interpretation von Fernerkundungsdaten auf den *Bildmaßstab* zu achten. Liegen mehrere Datensätze desselben Befliegungsgebiets vor, weisen diese nicht zwangsläufig denselben Maßstab auf. Auch innerhalb eines Datensatzes kann die geometrische Auflösung der Bilddaten variieren. So werden bei Aufnahmen der HRSC die Farbkanäle üblicherweise mit einer geringeren Auflösung aufgezeichnet als die panchromatischen Kanäle (s. Abschnitt 2.2). Um auch bei unterschiedlichen Bildauflösungen zu vergleichbaren Ergebnissen bei der Bildinterpretation zu gelangen, müssen sämtliche Größenangaben in Abhängigkeit von der Auflösung verwaltet werden. Die Größe von Filterkernen oder morphologischen Masken, Entfernungsangaben oder -schwellen, Flächenangaben usw. müssen in Maßen der Objektebene geführt, bzw. für Angaben in Pixeleinheiten entsprechend umgerechnet werden. Für einen Anwender sollten Angaben in Pixeleinheiten vermieden werden. Eingabemasken für die Angabe von Schwellwerten sollten stets auf Einheiten der Objektdomäne, wie Meter oder Quadratmeter, zurückgreifen.

### 1.5.3 Bildkanten als Objektgrenzen

Objekte wie Bäume oder Häuser lassen sich visuell gut voneinander unterscheiden. Auch die Differenzierung mehrerer Objekte desselben Typs bereitet der menschlichen Wahrnehmung im Allgemeinen keine Probleme. Für eine automatisierte Extraktion von Objekten kommen in der Bildverarbeitung die gleichen Erkenntnisse zum Tragen wie in der Wahrnehmungspsychologie. Ein Objekt erscheint als Einheit, wenn es sich in einem oder mehreren Attributen seiner Teile als gleichartig erweist. Ein fahrendes Auto wird durch die Gleichartigkeit der Bewegung aller Karosserieteile als *ein* Auto wahrgenommen. Ein Kleidungsstück zeichnet sich durch eine gemeinsame Farbe oder durch dieselbe Textur des Stoffes aus. In der Bildverarbeitung entspricht diese Art, Objekteinheiten zu bilden, der Segmentierung. Ziel der Segmentierung ist die Aufteilung eines Bildes in kleinere Einheiten, die gleichartig bezüglich eines zuvor definierten Attributs sind. Entsprechend dieser segmentbasierten Vorgehensweise werden in dieser Arbeit erhöhte Objekte bestimmt und Gebäude und Bäume unterschieden.

Im Unterschied zu einer Segmentierung, bei der homogene Bildpunkte zusammengefasst wer-

den, kann die Bildung von Einheiten auch anhand starker Unterschiede erfolgen. Diese Wechsel zeichnen üblicherweise die Objektgrenze aus, anhand derer die Unterscheidung benachbarter Objekte ermöglicht wird. Bei Betrachtung der Helligkeiten macht sich dieser Wechsel im Bild durch Grauwertsprünge, die Bildkanten, bemerkbar.

Viele Ansätze zur Kantendetektion beruhen auf der Annahme, dass schnelle Grauwertänderungen oder -diskontinuitäten durch Maxima des Bildgradienten oder Nulldurchgänge der zweiten Ableitung lokalisiert werden können. Der Gradient einer skalaren Funktion  $f$  ist definiert als

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} f_x \\ f_y \end{pmatrix}. \quad (1.1)$$

Bei digitalen Bilddaten ist zu berücksichtigen, dass Helligkeiten durch diskrete Grauwertstufen repräsentiert werden. Für eine Bildmatrix  $F$  werden die partiellen Ableitungen  $F_x$  und  $F_y$  entweder durch diskrete Differenzoperatoren oder durch Vergleich mit einer Reihe von Gradientenmasken („Template Matching“) annähernd bestimmt. Die Gradientenstärke  $|\nabla F|$  und die Gradientenorientierung  $\varphi(\nabla F)$  ergeben sich vektoriell als

$$|\nabla F| = \sqrt{F_x^2 + F_y^2}, \quad \varphi(\nabla F) = \arctan(F_x/F_y). \quad (1.2)$$

Aus Gründen der Rechenzeit wird oft einfach nur die Summe oder das Maximum der Absolutbeträge als Maß für die Gradientenstärke verwendet. Die Gradientenrichtung gibt die Richtung der größten Grauwertänderung an, der Betrag des Gradienten hingegen die Stärke der Änderung, also den Kontrast im betrachteten Punkt. Die meisten ableitungsorientierten Kantendetektoren identifizieren nun Kanten mit den Bildpunkten, an denen die Gradientenstärke einen Mindestwert aufweist und in einer Nachbarschaft von Pixeln gleicher Gradientenrichtung maximal ist.

Die in dieser Arbeit zum Einsatz kommenden Methoden nutzen den Bildgradienten allerdings direkt. Die Gradientenstärke und -orientierung ist die Basis für die Aufteilung der Gebäudesegmente in kleinere Einheiten sowie für die Nachbearbeitung der Konturlinie der Gebäude. Die Reduktion des Bildgradienten auf Kantenpixel ist für diese Aufgabenstellungen nicht angebracht, da auch auf Gradienteninformation in der Nähe der Kantenpixel zurückgegriffen wird. Wie in Abschnitt 1.3.1 geschildert, existieren in der Literatur jedoch zahlreiche andere Ansätze zur Gebäudedetektion, die auf vorab detektierten Bildkanten aufbauen. Das Einpassen von Liniensegmenten stellt einen weiteren Abstraktionsschritt dar. Auf die Vor- und Nachteile eines solchen Vorgehens wird ebenfalls in Abschnitt 1.3.1 eingegangen.



## 2 Der Multispektralscanner HRSC-A

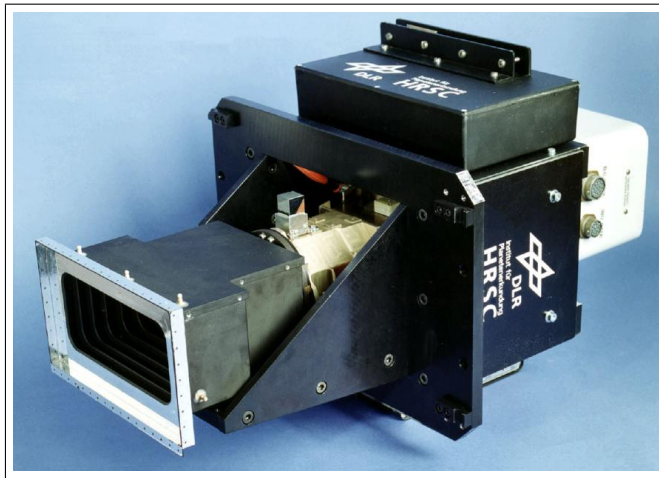
Wie in Abschnitt 1.2 geschildert, sind hochauflösende digitale multispektrale Orthobilder und digitale Oberflächenmodelle Datengrundlage dieser Arbeit. Sie stammen von der Erdfernerkundungskamera HRSC-A („*High Resolution Stereo Camera - Airborne*“) bzw. deren Folgemodellen. Prinzipiell sind die in den nachfolgenden Kapiteln vorgestellten Verfahren für eine automatisierte Objektextraktion nicht auf Daten dieses Sensors limitiert. Maßgebende Kriterien für den Einsatz anderer Standarddatenprodukte der Fernerkundung sind die Orthoperspektive, die Relation von Auflösung und Objektgröße sowie die Höheninformation.

Im Rahmen dieser Arbeit werden ausschließlich HRSC-Daten als Testdaten verwendet. Die Beschaffenheit der Daten spielt bei der Motivation konkreter Schritte für eine Bildinterpretation eine nicht unerhebliche Rolle. So ist beispielsweise der Schritt der Kantenadaption an Orthodaten (Kapitel 6) dadurch bedingt, dass aus einem HRSC-Höhenmodell Gebäudekanten nur relativ grob extrahiert werden können. Neben einem Einblick in eine relativ junge Sensortechnologie und deren Datenprozessierung stellt dieses Kapitel somit eine Diskussionsgrundlage für die folgenden Kapitel dar.

### 2.1 Aufbau und Eigenschaften

Die HRSC-A (s. Abb. 2.1(a)) ist ein multispektraler Zeilenscanner für fernerkundliche und photogrammetrische Fragestellungen. Charakteristisches technisches Merkmal der HRSC sind mehrere, parallel in der Fokalebene angeordnete *CCD-Zeilen* bei Nutzung einer gemeinsamen Optik. Auf die daraus resultierenden Blickwinkelunterschiede wird im nachfolgenden Abschnitt näher eingegangen. Während die räumliche Auflösung mit dem klassischen Luftbild vergleichbar ist (typisch sind 10–50 cm), heben sich die digitale Aufnahme der Daten und die vollautomatische Prozessierung bis hin zum orthorektifizierten Bild davon ab (s. Abschnitt 2.5). Die in der automatischen HRSC-Datenverarbeitung integrierte Orthorektifizierung ermöglicht die Korrektur von Verkippen am Rand des Flugstreifens. Bei Luftbildern ist dieser Korrekturschritt nur mit wesentlich höherem Aufwand möglich. Zu einem detaillierten Vergleich von Luftbild und HRSC sei auf die Arbeit von HOFFMANN (2001) verwiesen.

Ursprünglich wurde die HRSC für die russische Weltraummission Mars96 entwickelt (NEUKUM UND TARNOPOLSKY 1990, ALBERTZ ET AL. 1992, NEUKUM ET AL. 1995, HAUBER ET AL. 1996). Nach einem gescheiterten Missionsstart wurde das Qualifikationsmodell der HRSC zu einem flug-



(a) HRSC-A



(b) HRSC-AX

**Abbildung 2.1:** Die zwei Generationen der HRSC-A.

zeuggestützten Erdfernerkundungssensor modifiziert. Zur Dämpfung niedriger Vibrationsfrequenzen im Flugbetrieb wird die Kamera nun auf einer Stabilisierungsplattform montiert.

Im Lauf der letzten Jahre wurde die HRSC in Form neuer Kameramodelle fortentwickelt (das zusätzliche Kürzel „X“ wie „extended“ soll das verdeutlichen). Dabei konnte gezielter auf Anforderungen aus der Erdfernerkundung eingegangen werden. Bei Fertigstellung dieser Arbeit besteht die HRSC „Familie“ der neuen Generation aus einer Weitwinkelvariante HRSC-AXW002 sowie den prinzipiell baugleichen Modellen AX001, AX003 und AX004 (Abbildung 2.1(b) zeigt eine solche HRSC-AX). Wie Tabelle 2.1 zu entnehmen ist, wurden die Pixelanzahl und die maximale Auslesefrequenz mehr als verdoppelt sowie die Brennweite reduziert. Resultat ist eine deutlich höhere Bodenabdeckung pro Flugstreifen. Dadurch kann der Flugbetrieb wesentlich effizienter abgewickelt werden.

Techn. Parameter / HRSC-Typ	A	AX001	AXW002
Brennweite (mm)	175	151	47
Seitl. Öffnungswinkel, „FoV“ (°)	11,8	29	79
Anzahl der CCD-Zeilen	9	9	5
davon panchromatisch	5	5	3
Pixel pro Zeile	5 270	12 288	12 288
davon Bildpixel	5 184	12 000	12 000
Pixelgröße (μm)	7	6,5	6,5
Max. Auslesefrequenz (Hz)	450	820	1 640
Radiometr. Auflösung (Bit)	8	12	12

**Tabelle 2.1:** Technische Parameter unterschiedlicher HRSC-Modelle.

Zudem wurden die spektralen Charakteristika und die Blickwinkel der Kanäle geändert. Die ursprünglich für geologische Interpretationen des Mars ausgelegten Filter sind für Untersuchungen der Vegetation auf der Erde nur bedingt geeignet (auf diese Problematik wird im Rahmen der Vegetationssegmentierung in Abschnitt 4.2 näher eingegangen). Das für Vegetation typische starke Ansteigen der Reflektanz im Bereich um 700 nm (TUCKER 1979) kann durch die neuen, diesen Bereich direkt flankierenden roten und infraroten Kanäle nun gut erfasst werden. Auch die großen Blickwinkelunterschiede der HRSC-A Multispektralkanäle wurde verringert, um die Erfassung einer möglichst ähnlichen Oberflächenmorphologie sicherzustellen. In Tabelle 2.2 sind die spektralen Charakteristika und die Blickwinkel von HRSC-A, HRSC-AX001 und HRSC-AXW002 zusammengestellt.

In den nachfolgenden Abschnitten wird nun das Aufnahmeprinzip der HRSC erläutert und die Verarbeitung der HRSC-Daten von den Rohdaten bis hin zum Endprodukt Farb-Orthomosaik diskutiert. Diese Datenverarbeitungsschritte werden anschließend im Kontext einer automatisierten Objektinterpretation beleuchtet. In Abschnitt 2.4 wird dargestellt, in welchem Auflösungs- und Genauigkeitsbereich sich die HRSC-Datenprodukte befinden.

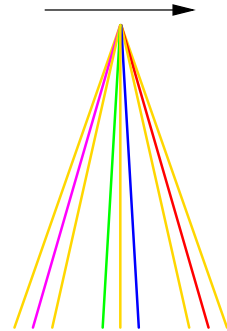
## 2.2 Das Aufnahmeprinzip

Die HRSC arbeitet nach dem sogenannten „Pushbroom“-Prinzip: Eine Bilddimension ergibt sich aus der Länge der CCD-Zeile, die andere aus der Flugbewegung (veranschaulicht in Abb. 2.2). Der Nadirkanal blickt (per Definition) senkrecht zur Bildebene. Durch die Anordnung der anderen Kanäle in der Fokalebene parallel zum Nadirkanal resultieren die in Tabelle 2.2 angegebenen Blickwinkel relativ zum Nadir. Diese Blickwinkelunterschiede sind Grundlage für die Berechnung eines Höhenmodells. Bei der Erzeugung multispektraler Farbkomposite sind die Blickwinkelunterschiede nicht erwünscht und müssen in einer Orthokorrektur photogrammetrisch entzerrt werden. Bei größeren Blickwinkelunterschieden und einem durch Höhengsprünge geprägten Terrain werden unterschiedliche Ansichten eines Objekts erfasst. Eine Zuordnung in dem Orthokorrekturschritt ist dann nicht möglich und führt zu entsprechenden Entzerrungsfehlern. Um diese Korrekturfehler zu minimieren wurden in der HRSC-AX001, wie bereits erwähnt, die Multispektralkanäle möglichs dicht um den Nadir gruppiert. Statt 15,9° (HRSC-A Rot-/Infrarotkanal) beträgt der maximale Blickwinkelunterschied zum Nadir jetzt 4,6° (HRSC-AX001 Blau-/Infrarotkanal) (s. dazu NEUKUM ET AL. 2000).

Neben den multispektralen Kanälen besitzen HRSC-A und HRSC-AX001 fünf panchromatische, denselben spektralen Bereich abdeckende Kanäle (Nadir, Stereo-, und Photometriekanäle). Sie werden zur Erzeugung eines Oberflächenmodells herangezogen. Die Photometriekanäle sollen zusätzlich einer (noch nicht implementierten) photometrischen Korrektur auf Basis der bidirektionalen Reflektanzverteilungsfunktion („*Bidirectional Reflectance Distribution Function*“, BRDF) dienen. Die richtungsabhängige Verteilung (Sonnenstand, Blickwinkel, Zusammensetzung und Neigung der Oberfläche) der reflektierten Energie soll dadurch normiert werden.

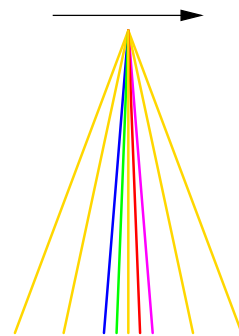
Kanalbezeichnung	Wellenl. (nm)	Blickw. (°)
Stereo vorne (panchrom.)	$675 \pm 90$	18,9
Rot	$750 \pm 20$	15,9
Photometrie vorne (panchrom.)	$675 \pm 90$	12,8
Blau	$440 \pm 45$	3,3
Nadir (panchrom.)	$675 \pm 90$	0,0
Grün	$530 \pm 45$	-3,3
Photometrie hinten (panchrom.)	$675 \pm 90$	-12,8
Infrarot	$970 \pm 45$	-15,9
Stereo hinten (panchrom.)	$675 \pm 90$	-18,9

(a) HRSC-A



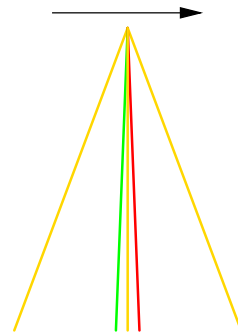
Kanalbezeichnung	Wellenl. (nm)	Blickw. (°)
Stereo vorne (panchrom.)	$640 \pm 120$	20,5
Photometrie vorne (panchrom.)	$640 \pm 120$	12,0
Infrarot	$792 \pm 22$	4,6
Rot	$662 \pm 20$	2,3
Nadir (panchrom.)	$640 \pm 120$	0,0
Grün	$553 \pm 23$	-2,3
Blau	$480 \pm 30$	-4,6
Photometrie hinten (panchrom.)	$640 \pm 120$	-12,0
Stereo hinten (panchrom.)	$640 \pm 120$	-20,5

(b) HRSC-AX001

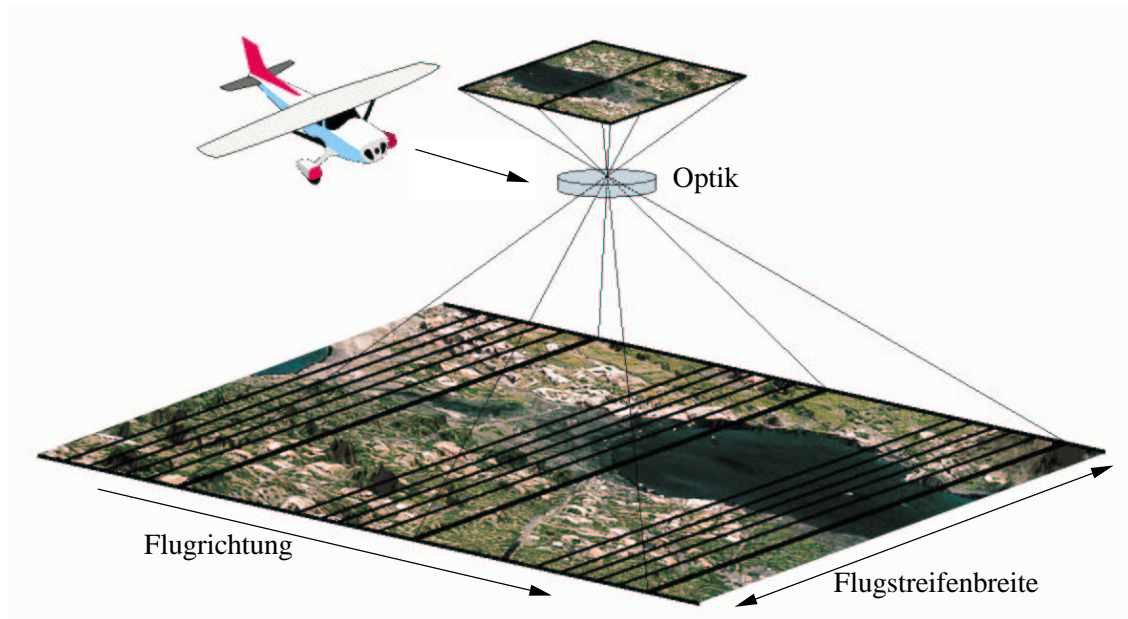


Kanalbezeichnung	Wellenl. (nm)	Blickw. (°)
Stereo vorne (panchrom.)	$632,5 \pm 117,5$	20,5
Rot	$625,0 \pm 55,0$	2,3
Nadir (panchrom.)	$632,5 \pm 117,5$	0,0
Grün	$530,0 \pm 45,0$	-2,3
Stereo hinten (panchrom.)	$632,5 \pm 117,5$	-20,5

(c) HRSC-AXW002


**Tabelle 2.2:** Spektrale Empfindlichkeit und Blickwinkel der Kanäle unterschiedlicher HRSC-Modelle.





**Abbildung 2.2:** Das Pushbroom-Scanprinzip der HRSC.

Um je nach Beleuchtung eine optimale radiometrische Aussteuerung zu gewährleisten, kann während des Flugs die Signalverstärkung variiert werden. Nach Sichtung eines Bildvorprodukts während des Flugs („Online Quicklook“) kann aus einer Reihe von Verstärkungswerten gewählt werden. Diese bestimmen die Dämpfung des Signals innerhalb der CCD wie auch die Verstärkung durch die nachfolgende Elektronik (vor der A/D-Wandlung).

Die Auslesefrequenz der CCD-Zeilen ist in Stufen bis zu einer Maximalfrequenz (s. Tabelle 2.1) frei wählbar. Um die maximale Auslesefrequenz zu erhöhen, werden zwei (gerade und ungerade Pixeladressen werden getrennt ausgelesen) bzw. vier Ausleseelektroniken (zusätzliche Aufteilung in linke und rechte Zeilenhälfte, nur HRSC-AXW002) pro CCD-Zeile verwendet. Eine Änderung der Auslesefrequenz ist gekoppelt an eine inverse Änderung der Integrationszeit. Je länger ein CCD-Element belichtet wird, desto höher fällt der Signal-zu-Rausch-Verhältnis aus. Neben der radiometrischen Güte beeinflusst die Auslesefrequenz aber auch die geometrische Auflösung in Flugrichtung. Vereinfacht ist die Auflösung in Flugrichtung,  $s_y$ , unabhängig von Flughöhe und Brennweite<sup>1</sup>. Sie ergibt sich aus dem Gesetz zur gleichförmigen Bewegung:

$$s_y = v \cdot t_{\text{int}} , \quad (2.1)$$

dabei geben  $v$  die Fluggeschwindigkeit und  $t_{\text{int}}$  die Integrationszeit an. Die Auflösung quer zur Flug-

<sup>1</sup>Dabei wird vernachlässigt, dass sich der Überlappungsanteil aufeinanderfolgender Pixel verändert. Zu Überlappungen kommt es wegen einer Überabtastung sowie der durch die Optik bedingten flächenhaften Verteilung eines Objektpunkts auf der Bildebene.

richtung,  $s_x$ , ist unabhängig von der Fluggeschwindigkeit und ergibt sich nach den Strahlensätzen als

$$s_x = \frac{s_p \cdot h}{f}, \quad (2.2)$$

wobei  $s_p$  der Pixelabstand auf der CCD-Zeile,  $h$  die Flughöhe und  $f$  die Brennweite des Kamerasystems darstellen. Je nach Aufnahmekonstellation können sich  $s_x$  und  $s_y$  also durchaus unterscheiden; quadratische Pixel sind bei den Rohdaten die Ausnahme. Für die HRSC-A ergibt sich bei einer Integrationszeit von 2,24 ms sowie einer typischen Flughöhe von 3 000 m und -geschwindigkeit von 70 m/s somit eine Pixelabdeckung von  $12 \times 15,68$  cm.

Zusammen mit den Bilddaten werden über ein D-GPS-System („Differential Global Positioning System“) die aktuelle Position und über ein inertiales Messsystem (INS, Inertial Navigation System) die Lage im Raum erfasst. Diese direkte Georeferenzierung jeder Bildzeile ist Grundlage für die spätere Korrektur der Flugbewegung in den Bilddaten (s. Abschnitt 2.3.3).

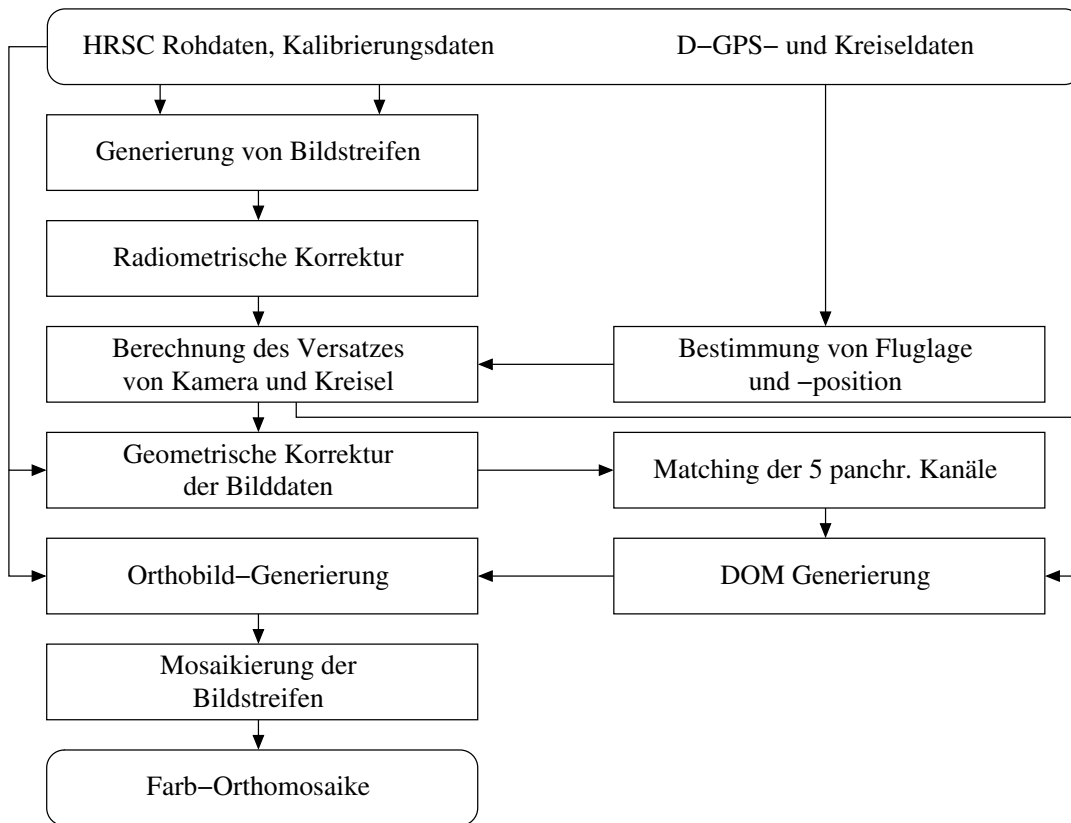
Die Speicherung der Daten im Flugzeug erfolgt momentan ausschließlich über ein Bandlaufwerk<sup>2</sup>. Die Bandbreite der Datenübertragung ist dadurch begrenzt auf 32 MB/s. Bei der HRSC-A spielt diese Grenze aufgrund einer anderen Beschränkung keine Rolle: Die Kameraelektronik kann nur 9,5 MB/s verarbeiten. In der AX-Generation entfällt diese Grenze, jedoch könnte die Bandbreite des Bandlaufwerks um ein Vielfaches übertroffen werden. So resultiert beim Auslesen aller HRSC-AX003 Kanäle mit einer Integrationszeit von 1,22 ms und 12 Bit Auflösung eine Datenrate von 128,45 MB/s. Neben einer Verlängerung der Integrationszeit für ausgewählte Kanäle wurde deshalb die Möglichkeit einer dazu äquivalenten räumlichen Integration der Bildinformation geschaffen. Dazu wird die Information von zwei, vier oder acht benachbarten Pixeln zu einem sogenannten Makropixel gemittelt. Ferner besteht in der AX-Generation die Möglichkeit, für einzelne Kanäle die radiometrische Auflösung von 12 Bit noch innerhalb der Kameraelektronik auf 8 Bit zu reduzieren. Dies geschieht entweder durch Verzicht auf die vier niedrigsten Bits oder durch Anwendung einer zuvor definierten Umwandlungstabelle. Eine typische Befliegungskonfiguration für die HRSC-AX003 sieht folgendermaßen aus:

Kanal	$t_{\text{int}}$ (ms)	Makropixel	Datenformat (Bit)
Nadir	2,844	1	12
Stereo	2,844	2	12
Photometrie	5,688	2	12
Farbkanäle	5,688	2	12

Daraus resultiert eine Datenübertragungsrate von 24,49 MB/s.

---

<sup>2</sup>Alternativ sollte auch ein Plattensystem mit SCSI-Schnittstelle erprobt werden.



**Abbildung 2.3:** Darstellung der HRSC-Datenverarbeitung.

## 2.3 Die HRSC-Datenverarbeitung

Die während des Flugs auf ein Band gespeicherten Bilddaten werden nahezu vollautomatisiert am Verarbeitungsstandort verarbeitet. Der Ablauf der Verarbeitungskette ist in Abbildung 2.3 schematisch dargestellt und wird in den folgenden Abschnitten näher betrachtet. Zur Veranschaulichung zeigt Abbildung 2.4 Bildbeispiele zu den jeweiligen Verarbeitungstufen. Grundlage dafür ist der Grün-Kanal einer Berlin-Befliegung mit der HRSC-AX004. Die Farbkanäle weisen eine geometrische Auflösung von 30 cm auf.

### 2.3.1 Generierung der Bildstreifen

Auf dem Magnetband werden die Bilddaten sequenziell (Kanal nach Kanal, Zeile für Zeile) in Form einer sogenannten „Multiline“-Datei abgelegt. Für jeden Flugstreifen wird die Aufnahme typischerweise an- und wieder abgeschaltet, so dass daraus eine Bilddatei resultiert. Nachdem die Daten an der Verarbeitungsstation vom Band eingelesen wurden, können die Multiline-Dateien wieder in die einzelnen Kanäle zerlegt werden („Demultiplexen“). Daten mit einer radiometrischen Auflösung von 12-Bit werden als 16-Bit Ganzzahlwerte gespeichert. Ergebnis dieses als „systematische Verar-

beutung“ bezeichneten Vorgangs sind Bilddateien für jeden Flugstreifen und jeden aufgenommenen Kanal. Jede Datei beginnt mit einem Dateikopf, in dem die Aufnahmeparameter sowie die bereits erfolgten Verarbeitungsschritte festgehalten sind. Der linke obere Teil von Abbildung 2.4 zeigt ein Bildbeispiel der resultierenden Rohdaten.

### 2.3.2 Radiometrische Korrektur

Anhand der radiometrischen Korrektur sollen die Grauwerte verschiedener Pixel vergleichbar gemacht werden. Die durch die Kameraelektronik und -optik bedingten radiometrischen Unterschiede werden dabei weitestgehend behoben.

Das Grundrauschen der CCD-Elemente wird in der *Dunkelstromkorrektur* vom Signal der belichteten Pixel abgezogen. Da das Dunkelstromrauschen nicht konstant ist, gibt es auf jeder CCD-Zeile unbelichtete Pixel, die nur der Erfassung dieses Dunkelstroms dienen.

Betrachtet man den Strahlungsfluss durch ein abbildendes Objektiv, ergibt sich das auf einer Sensorfläche gemessene Signal  $S$  (Zahl der Elektronen) als

$$S = t_{\text{int}} \frac{\pi}{4} \left( \frac{D}{f} \right)^2 A_{\text{CCD}} \cos^4 \phi \int R_{\text{CCD}}(\lambda) \tau(\lambda) L(\lambda) d\lambda, \quad (2.3)$$

wobei  $t_{\text{int}}$  die Integrationszeit,  $\frac{D}{f}$  das Öffnungsverhältnis,  $A_{\text{CCD}}$  die Fläche eines Pixels,  $\phi$  den Blickwinkel in Bezug zum Nadir,  $R_{\text{CCD}}$  die spektrale Empfindlichkeit des Sensors,  $\tau$  die Transmission von Optik und Spektralfiltern sowie  $L$  die Strahldichte vor dem Sensor darstellen. Die Kameraoptik bedingt also einen zu  $\cos^4 \phi$  proportionalen Abfall des Signals zum Rand der CCD-Zeile hin.

Ferner ist es selbst unter Reinraum-Bedingungen nicht möglich, eine vollkommen saubere CCD-Oberfläche zu garantieren. So führen Verunreinigungen dazu, dass einzelne Pixel weniger Lichtenergie erhalten als der Rest. Beide Effekte sind in entsprechenden Bildausschnitten der in Abbildung 2.4 gezeigten Rohdaten zu erkennen. In der Vergrößerung ist die geringere Empfindlichkeit eines Pixels, in der Verkleinerung auf Flugstreifenbreite das Absinken des Signals zum Rand der CCD-Zeile hin zu erkennen.

Eine weitere Erschwernis für die Vergleichbarkeit der Pixelgrauwerte ist die bereits in Abschnitt 2.2 erwähnte Beschleunigung des Auslesevorgangs durch mehrere Ausleseelektroniken. Jede elektronische Schaltung weist ein leicht unterschiedliches Verstärkungsverhalten auf. Da gerade und ungerade Pixeladressen immer durch zwei verschiedene elektronische Schaltungen ausgelesen werden, kann die bildhafte Aneinanderreihung der Pixel zu Spaltenmustern führen. Zur Korrektur dieser Effekte, also einer *Normierung der Pixelempfindlichkeit*, werden im Rahmen der Kamerakalibrierung Aufnahmen vor einer Kugel mit homogener Lichtverteilung („Ulbrichtkugel“) durchgeführt. Daraus werden für jedes Pixel Empfindlichkeitsfaktoren abgeleitet, die dann zur Korrektur realer Bilddaten herangezogen werden. Werden Veränderungen an der Empfindlichkeit der CCD-Elemente festgestellt, müssen neue Ulbrichtkugelmessungen erfolgen. In Abbildung 2.4 wird das Resultat der

radiometrischen Korrektur durch eine Gegenüberstellung mit den Rohdaten besonders deutlich.

Weiterhin dienen die Messungen an der Ulbrichtkugel („Flatfield-Messung“) dazu, die Kanäle relativ zueinander zu kalibrieren. Ein Intensitätsanstieg des einfallenden Lichts soll zu ähnlichen Grauwertanstiegen in allen Kanälen führen — unabhängig von der gewählten Verstärkungsstufe. Hervorzuheben ist jedoch, dass die Kanäle *nicht* absolut radiometrisch kalibriert sind. Die gleiche Lichtintensität kann aufgrund der unterschiedlichen Verstärkungsfaktoren zu verschiedenen Grauwerten führen. Somit ist die HRSC nicht zu absoluten radiometrischen Messungen in Bilddaten geeignet.

Für nachfolgende Verarbeitungsschritte (z. B. Mosaikierung von Flugstreifen) werden innerhalb der radiometrischen Korrektur auch statistische Parameter berechnet und abgelegt. So werden Grauerthistogramme und Angaben über die Dynamik für jeden Flugstreifen und Kanal ermittelt.

### 2.3.3 Geometrische Korrektur

Bei der geometrischen Kalibrierung der Kamera wird die Lage der CCD-Zeilen relativ zu einem Kamerafixpunkt vermessen (Bestimmung der inneren Orientierung, s. z. B. SCHENK (1999, Kapitel 10.4)). Durch das kombinierte GPS/Inertialsystem werden die aktuellen Lagedaten aber in Bezug auf die Achsen des Inertialsystems ermittelt. Da sich durch jeden Einbau der Kamera ins Flugzeug der Abstand von Inertialsystem und Bildebene leicht ändert, muss er anhand identischer Punkte in überlappenden Flugstreifen bei jedem Flug erneut photogrammetrisch bestimmt werden („Bore-sight Alignment“). Ist dies geschehen, kann während eines Bildflugs mittels GPS und Inertialsystem die momentane Position und Lage für jede CCD-Zeile ermittelt werden. Da die Aufnahmetaktrate im Allgemeinen höher ist als die GPS Taktrate (200 Hz), müssen die aus dem GPS ermittelten Positionsdaten entsprechend interpoliert werden. Dann können jeder Bildzeile die Position und Lage der zugehörigen CCD zugeordnet werden („direkte Georeferenzierung“)<sup>3</sup>. Diese Navigationsdaten werden zusammen mit der Bildinformation gespeichert.

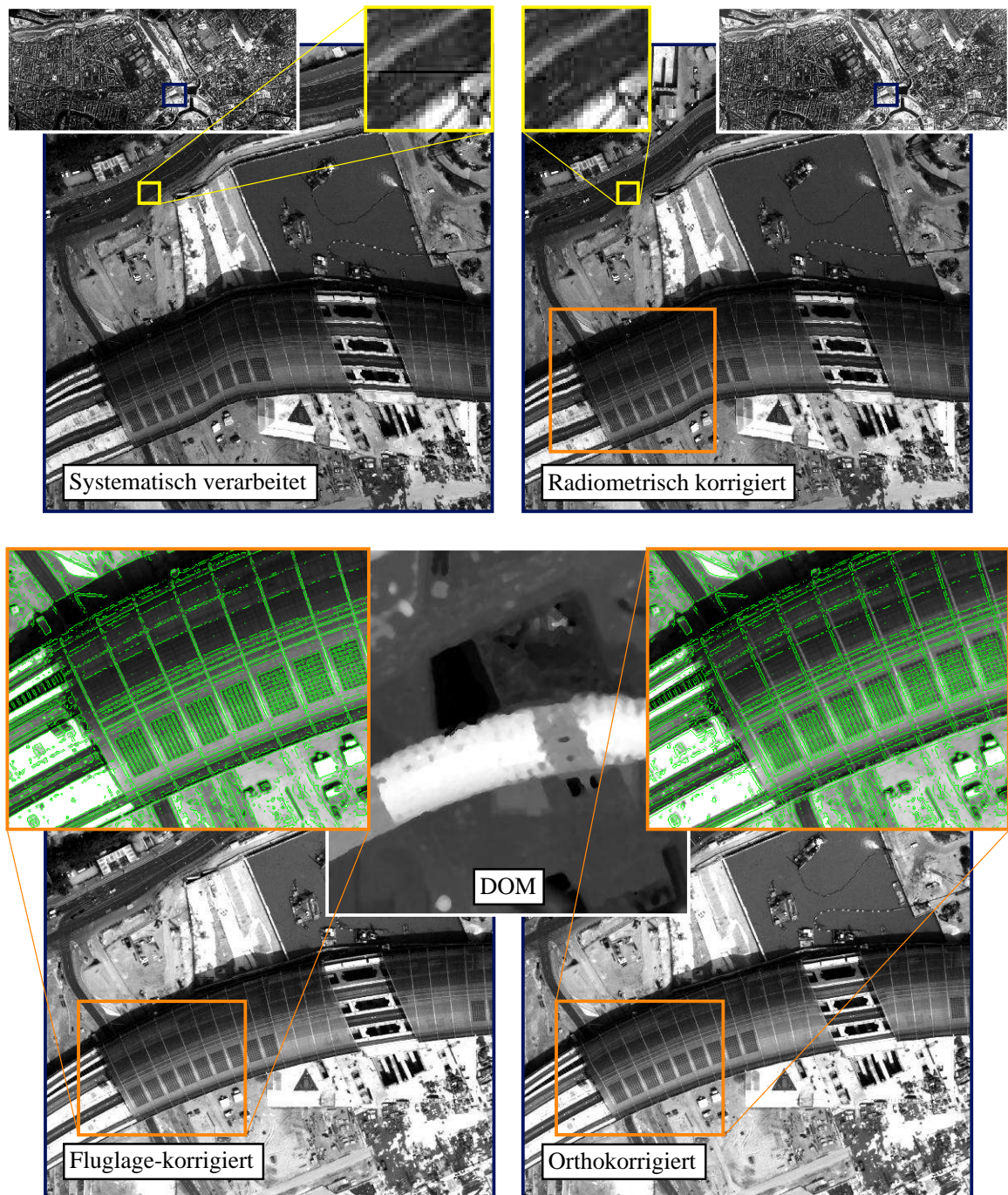
Während eines Bildflugs ändert sich die äußere Orientierung (Lage des Projektionszentrums und die Richtung der Aufnahmeachse, s. z. B. SCHENK (1999, Kapitel 10.4)) ständig. Aufgrund der gespeicherten Positions- und Lageinformation ist es in der geometrischen Korrektur (auch Fluglage- oder Vorkorrektur) nun möglich, die Bilddaten bezüglich der Fluglage zu entzerren. Ziel dieses Korrekturschritts ist, dass bei den Stereokanälen Parallaxen nur noch an Höhenunterschieden in Flugrichtung auftreten.

Entsprechend der Abweichung der Kameraachse von der Senkrechten werden die Daten einer Bildzeile verschoben und gestaucht bzw. gedehnt. Das Ergebnis stellt eine Rektifizierung auf eine mittlere Geländehöhe dar. Ist bereits ein grobes Geländemodell der Szene vorhanden, so kann die Entzerrung auch diesbezüglich erfolgen.

Die Auswirkungen der geometrischen Korrektur zeigen sich in Abbildung 2.4 bei einem Vergleich

---

<sup>3</sup>Da die Geometrie der CCD-Zeilen auf der Bildebene fix ist, gilt dies sogar für jedes Pixel.



**Abbildung 2.4:** Bildbeispiele zur photogrammetrischen Verarbeitung (Anmerkungen dazu im Text).



mit den radiometrisch korrigierten Bilddaten insbesondere durch Betrachtung regulärer Objektgeometrien, wie Linien und Linienkonstellationen.

Die resultierenden Bilddaten der HRSC-Kanäle sind bis auf den Einfluss der Abweichungen von der angenommenen Geländehöhe deckungsgleich (WEWEL UND SCHOLTEN 2000). Höhenunterschiede stellen sich nun in Form von Parallaxen dar. Sie sind die Basis für die im nächsten Abschnitt behandelte Berechnung des Oberflächenmodells.

### 2.3.4 Generierung eines Oberflächenmodells

Durch die Zuordnung/Korrelation („Image Matching“) der geometrisch vorkorrigierten Stereobildaten werden Objekthöhen bezüglich der mittleren Geländehöhe (s. geometrische Korrektur) bestimmt. Da die Bildzuordnung separat für jeden Flugstreifen erfolgt, der Flugweg dafür nahezu geradlinig ist und die HRSC-Öffnungswinkel relativ gering sind, sind die Parallaxen quer zur Flugrichtung sehr gering ( $< \pm 1$  Pixel, WEWEL UND SCHOLTEN (2000)). Dadurch kann der Suchbereich für die Bildzuordnung erheblich eingeschränkt werden.

Die Bestimmung *homologer* Bildpunkte (gleiche Objektpunkte in den unterschiedlichen Stereokanälen) erfolgt durch eine flächenhafte Grauwert-Korrelation (s. z. B. SCHENK 1999, Kapitel 10.4). Dabei wird der Nadirkanal als Referenzbild definiert; alle anderen Stereokanäle werden diesem Referenzbild gleichzeitig zugeordnet. Die Zuordnung erfolgt in Form eines hierarchischen Pyramidenverfahrens (WEWEL 1996). Ausgehend von den Pyramidenstufen mit der gröberen Auflösung werden identische Punkte durch Kreuzkorrelation bestimmt und schrittweise bis zur feinsten Stufe übertragen. Dort werden die endgültigen homologen Bildpunkte durch Korrelation anhand des kleinsten quadratischen Fehlers mit Genauigkeiten im Bereich von etwa einem halben Pixel bestimmt.

Je nach Übereinstimmungsgrad werden durch die homologen Punkte und die Koordinaten der Originaldaten<sup>4</sup> bis zu fünf räumliche Sichtstrahlen definiert. Wegen der fehlenden Redundanz werden homologe Punkte mit weniger als drei Sichtstrahlen von den weiteren Berechnungen ausgeschlossen. Die fehlende Zuordnung eines Punktes in den Stereokanälen kann verschiedene Ursachen haben: In direkter Nachbarschaft von großen Höhengsprüngen befinden sich Objektpunkte im Schattenschatten eines Teils der Stereokanäle. Aus dem kanalspezifischen Blickwinkel verdeckt das hohe Objekt also diese Punkte. Ferner können sich bei der flächenhaften Zuordnung der Bilddaten Schwierigkeiten ergeben. Die Zuordnung linienhafter Elemente wie z. B. Schienensträngen kann zu einem verbleibenden Freiheitsgrad führen. Die Zuordnung wird ebenfalls misslingen, wenn die Zuordnungsausschnitte mangelhaft texturiert sind, wie dies z. B. bei einer Wasseroberfläche der Fall ist.

Die aus dem Schnitt der Sichtstrahlen resultierenden Schnittpunkte werden durch Ausgleichung zu einem Objektpunkt reduziert. Die kartesischen Objektpunktkoordinaten werden dann entsprechend der gewünschten Kartenprojektion transformiert. Da nicht allen Rasterpunkten des DOM durch den Schnitt der Sichtstrahlen ein Höhenwert zugeordnet wird, werden die verbleibenden Punkte mittels

---

<sup>4</sup>Eine Verknüpfung zu den Originaldaten bleibt in den geometrisch korrigierten Daten erhalten.

Interpolation gefüllt. Je nach Anwendung kommen dazu Verfahren zum Einsatz, die dem Mittelwert- oder dem Medianfilter vergleichbar sind. Die Daten werden in der Regel als 16-Bit-Ganzzahlenwerte abgespeichert. In der üblichen Dezimeterabstufung ergibt sich so ein Höhenbereich von  $\pm 3\,276\text{ m}$  bezüglich einer wählbaren mittleren Höhe. Abbildung 2.4 zeigt in der Mitte des unteren Teils das zu dem Bildbeispiel gehörige DOM.

### 2.3.5 Orthobild-Generierung

Die Erstellung eines Orthobildes, also die Simulation der Parallelperspektive, erfolgt nach SCHOLTEN (1996) analog zur geometrischen Korrektur der Bilddaten. Mit Hilfe der bekannten Orientierung und der nun vorliegenden Oberflächenform werden die radiometrisch korrigierten Bilddaten projiziert. Dazu wird der für jeden Oberflächenpunkt definierte Sichtstrahl mit der durch das DOM beschriebenen Geländefläche geschnitten (Strahlverfolgung, „Ray-Tracing“). Aus Effizienzgründen geschieht dies nicht für alle Bildpunkte, sondern nur im Abstand der meist höheren DOM-Rasterweite (sog. Ankerpunkte, s. HILDEBRANDT (1996, Abschnitt 4.5.9.3)). Für die verbleibenden Bildpunkte wird die Transformation anhand der vier sie umgebenden, bereits projizierten Ankerpunkte ermittelt. Ihr Grauwert wird durch bilineare Interpolation der radiometrisch korrigierten Bilddaten abgeleitet<sup>5</sup>. Auf diese Art wird für sämtliche Spektralkanäle und Flugstreifen ein Orthobild erzeugt.

Da der für das Bildbeispiel in Abbildung 2.4 gewählte Grünkanal nur eine geringe Blickwinkel-differenz zum Nadir aufweist, ist ein Unterschied von geometrisch korrigierten zu orthokorrigierten Bilddaten nur schwer auszumachen. Deswegen werden mit einem Kantendetektor die Bildkanten in den geometrisch korrigierten Daten extrahiert. Sie sind den Vergrößerungen der beiden Korrekturstufen grün überlagert. Nun wird sichtbar, dass die Bilddaten an Höhenunterschieden entsprechend verschoben sind. Auf ebener Fläche hingegen sind die Bildkanten unverändert.

### 2.3.6 Mosaik-Erstellung

Bei der Bearbeitung größerer Projektgebiete, die mehr als einen Flugstreifen umfassen, erfolgt zuletzt eine Mosaikbildung der einzelnen Orthobildstreifen. Hierbei werden in einem iterativen Verfahren die Grauwertistogramme der einzelnen Streifen aneinander angepasst. Anhand der mehrfach vorhandenen Bildinformation in sich überlappenden Streifen werden entsprechende Tabellen zur Korrektur der Histogramme berechnet (WEWEL ET AL. 1999). Sie sind die Grundlage für einen gleitenden Übergangs der sich überlappenden Bildstreifen, soweit die Form der Überlappung dies zulässt. Voraussetzung für das Zustandekommen eines homogen erscheinenden Bildmosaiks ist eine konstante Beleuchtungssituation *innerhalb* eines Flugstreifens. Veränderungen der Helligkeit in einem Flugstreifen werden nicht korrigiert. Die so beschriebene Mosaik-Erstellung erfolgt separat für alle Kanäle.

---

<sup>5</sup>Auf Wunsch können die Lücken auch durch den nächsten Nachbarn oder bikubische Interpolation gefüllt werden. Das erstgenannte Verfahren kommt vorzugsweise dann zum Einsatz, wenn eine multispektrale Klassifikation erfolgen soll.



## 2.4 Geometrische Qualität

Um die geometrische Qualität der HRSC-Bilddaten beurteilen zu können, muss ein Test des Gesamtsystems unter realen Bedingungen erfolgen. Die obligatorische Vermessung der Kamera im Kalibrierungslabor führt zwar ebenfalls zu Angaben über Punktlagegenauigkeiten, doch fehlen hier weitere „Fehlerquellen“ wie Schwingungen im Flugzeug oder Ungenauigkeit des GPS-/Inertialsystems. Deshalb wird mit dem Kamerasystem ein Testfeld befliegungsüblich aufgenommen. Im Testfeld sind zahlreiche Kontrollpunkte markiert. Sie sind entweder direkt per GPS vermessen oder aber durch Aerotriangulation aus letztgenannten Punkten bestimmt. Für die HRSC-A wurde 1998 eine Befliegung in 3 000 m Flughöhe bei durchschnittlich 72 m/s durchgeführt (WEWEL ET AL. 1998, WEWEL UND BRAND 1999, SCHOLTEN UND WEWEL 2000). Der Nadirsensor wurde mit einer Integrationszeit von 2,24 ms und die Stereosensoren mit 4,48 ms ohne Pixelzusammenfassung aufgezeichnet. Die mittlere Abweichung zu den Kontrollpunkten ergibt sich in dieser Testbefliegung zu 12 cm in der Ebene und 17 cm in der Höhe (0,04 ‰ bzw. 0,06 ‰ der Flughöhe). Daraus resultiert eine Punktgenauigkeit von 21 cm (0,07 ‰ der Flughöhe). Selbst für die Nadirpixelauflösung von  $12 \times 16$  cm liegt die mittlere Ungenauigkeit somit immernoch im Bereich eines Pixels.

Eine ähnliche Befliegung wurde im Jahr 2000 für die HRSC-AXW002 durchgeführt (s. SCHOLTEN ET AL. 2001). Für diesen Test wurde in 1 600 m Höhe mit etwa 85 m/s geflogen. Die panchromatischen Sensoren wurden mit 2,24 ms ohne Pixelzusammenfassung aufgezeichnet. Bei einer Pixelabmessung von  $21 \times 20$  cm resultiert eine mittlere Punktgenauigkeit von 25 cm (0,16 ‰ der Flughöhe) aus den Testdaten. Für die HRSC-AX001 fand im Jahr 2001 eine Testbefliegung in 1 000 m Höhe mit etwa 70 m/s statt (s. ebenfalls SCHOLTEN ET AL. 2001). Die panchromatischen Sensoren wurden mit 2,24 ms ohne Pixelzusammenfassung aufgezeichnet. Bei einer Pixelabmessung von  $4 \times 16$  cm liegt die resultierende mittlere Punktgenauigkeit bei 12 cm (0,12 ‰ der Flughöhe). Somit bewegt sich die mittlere geometrische Genauigkeit bei allen HRSC-Modellen im Bereich eines Pixels.

## 2.5 Die automatisierte Interpretation von HRSC-Daten

Werden HRSC-Daten für eine automatische Bildinterpretation verwendet, ist ein möglicher Einfluss der im letzten Abschnitt erläuterten Schritte zur HRSC-Datenverarbeitung vorab einzubeziehen. In diesem Abschnitt werden die für die nachfolgenden Schritte relevanten Aspekte herausgearbeitet. Zum einen werden in Abschnitt 2.5.1 radiometrische Kriterien beleuchtet, zum anderen in Abschnitt 2.5.2 Kriterien, die ein HRSC-Oberflächenmodell betreffen.

### 2.5.1 Radiometrie

Die großen Datenmengen in der Erdfernerkundung sind Hauptmotivation für eine Automatisierung der Interpretation von Bilddaten. Wird ein großes Zielgebiet über mehrere Stunden oder gar Tage hin-

weg aufgenommen, so ist mit sich deutlich ändernden Beleuchtungsbedingungen zu rechnen. Ferner beeinflussen Lichtstreuungen die Radiometrie des Bildes. Sowohl atmosphärische wie auch objektspezifische Streuungen sind abhängig von der Wellenlänge und vom Blickwinkel des einfallenden Lichts (HILDEBRANDT 1996, s. z. B.). Sie wirken sich daher unterschiedlich stark auf die einzelnen HRSC-Kanäle aus.

Wie in Abschnitt 2.3.2 geschildert, sind die HRSC-Kanäle nur relativ zueinander kalibriert. Eine absolute radiometrische Kalibrierung, wie sie zur Korrektur unterschiedlicher Beleuchtungsbedingungen wie auch zur Atmosphärenkorrektur nötig wäre, steht noch aus. Bei einer Mosaikierung von Flugstreifen (s. Abschnitt 2.3.6) wird zwar versucht, die Streifen histogrammbasiert radiometrisch auszugleichen, jedoch steht dabei ein für das Auge gefälligeres Produkt im Vordergrund. Beleuchtungsparameter wie Flugzeit und Sonnenstand wurden erst jüngst in die Prozessierung integriert. Dieser Hintergrund erschwert eine multispektrale Auswertung der Bilddaten. Ein Grauwert von 120 kann in einem anderen Flugstreifen durchaus einem Grauwert von 80 entsprechen. Einfache multispektrale Klassifikationsverfahren werden durch Beleuchtungsabhängigkeit beeinträchtigt.

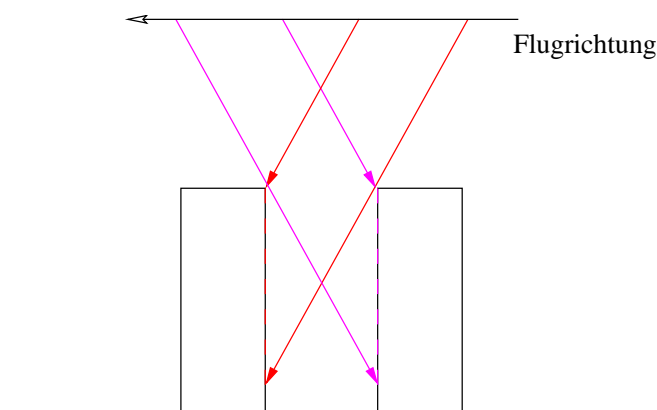
Daher bietet es sich an, nicht die Grauwerte selbst, sondern deren Beziehung zueinander zu klassifizieren. So geben Texturmerkmale die räumliche Verteilung der Grauwerte in einer bestimmten Nachbarschaft an. Um bei der Texturklassifikation intensitätsunabhängig zu werden, muss auf den ursprünglichen Grauwert als Klassifikationsparameter verzichtet werden.

Bei multispektralen Daten gibt es zudem die Möglichkeit, die Kanäle einander gegenüber zu stellen. Ratios aus multispektralen Kanälen sollten wesentlich beleuchtungsunabhängiger sein als die Daten der Kanäle selbst.

### **Schatten und Sichtschatten**

Insbesondere bei einem großen Projektgebiet und damit einer langen Befliegungsdauer weisen die Bildstreifen aufgrund des sich ändernden Sonnenstandes unterschiedliche Helligkeiten auf. Wird daraus ein Mosaik berechnet, so ist zwar eine Angleichung der Radiometrie nach Abschnitt 2.3.6 möglich. Der gleitende Übergang im Überlappungsbereich ist jedoch nur gerechtfertigt, wenn sich dort während der Zeitspanne zwischen den beiden Überflügen keine inhaltlichen Änderungen ergeben haben. Doch gerade wenn ein Bildstreifen (z. B. bei schlechtem Wetter) nachträglich beflogen wird, werden die Schatten durch den geänderten Sonnenstand unterschiedlich stark ausgeprägt sein.

Durch die unterschiedlichen Blickwinkel der HRSC werden Oberflächenpunkte in jedem Kanal in einer anderen Ansicht erfasst. Bei Höhengsprüngen, wie sie in der Stadt üblich sind, ist ein Teil der Bodenpunkte durch die erhöhten Objekte für bestimmte Blickwinkel verdeckt. Diese Bodenpunkte sind dann für Kanäle mit diesen Blickwinkeln „unsichtbar“, sie befinden sich im Sichtschatten der jeweiligen Kanäle. Abbildung 2.5 verdeutlicht diesen Zusammenhang. Hier wird der Boden zwischen den beiden Gebäuden weder von dem vorwärts noch von dem rückwärts blickenden Kanal erfasst. Stattdessen wird ein Teil der Gebäudefassade aufgenommen. Da bei der Orthokorrektur der



**Abbildung 2.5:** Schematische Darstellung zur Blickwinkelproblematik. Der Grund zwischen den beiden Häusern kann von den schrägblickenden Kanälen (magenta und rot dargestellt) nicht erfasst werden.

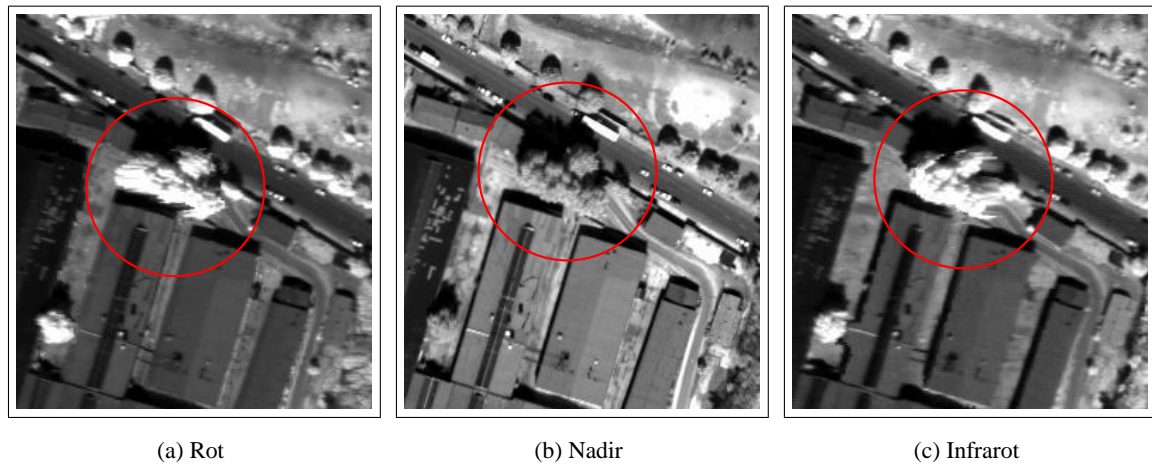
senkrechte Blickwinkel von oben simuliert wird, fehlt in den schrägblickenden Kanälen bei Höhengsprüngen die dafür notwendige Bildinformation. Nach Abschnitt 2.3.5 werden diese Lücken durch Interpolation der Nachbarwerte gefüllt. Die Sichtbarkeit von Bodenpunkten und somit der Interpolationsumfang ist dadurch in allen Kanälen verschieden. Bei HRSC-Orthobildern ist eine multispektrale Interpretation (z. B. Klassifikation) von Grundpunkten in der Nachbarschaft (in Flugrichtung) von Höhengsprüngen somit nicht sinnvoll. Diese Problematik wird in Abbildung 2.6 durch ein Bildbeispiel veranschaulicht. In den drei Orthobildern aus HRSC-A<sup>6</sup> Rot-, Nadir- und Infrarotdaten ist eine höhere Baumgruppe im Bildmittelpunkt. Die Flugrichtung geht von Ost nach West, hier also von rechts nach links. Bei Rot- und Infrarotkanal, die einen Blickwinkelunterschied von  $31,8^\circ$  haben, ist die Interpolation in Flugrichtung deutlich zu erkennen.

### Zeitversetzte Aufnahme

Aus den unterschiedlichen Blickwinkeln resultieren auch unterschiedliche Zeitpunkte an denen ein Bodenpunkt durch die Kanäle erfasst wird. So besteht bei einer HRSC-A Aufnahme bei einer typischen Flughöhe von 3 000 m mit einer Geschwindigkeit von 70 m/s zwischen den beiden äußeren Stereokanälen ein Zeitunterschied von 29,4 s und zwischen den beiden äußeren Farbkanälen (Rot und Infrarot) einen Unterschied von 24,4 s. Falls in dieser Zeitspanne Veränderungen im Aufnahmebereich stattfinden, hat dies Auswirkungen auf ein multispektrales Bildprodukt. Da Veränderungen der Beleuchtungssituation innerhalb eines Flugstreifens weniger wahrscheinlich sind, werden nachfolgend nur Veränderungen durch sich bewegende Objekte betrachtet<sup>7</sup>.

<sup>6</sup>Die HRSC-A weist den größten Blickwinkelunterschied zwischend den Farbkanälen auf; daher sind die Daten besonders gut für das Bildbeispiel geeignet.

<sup>7</sup>Eine unterschiedliche Beleuchtungssituation benachbarter Flugstreifen kann jedoch die Qualität des Bildmosaiks beeinträchtigen. Dies wurde bereits in Abschnitt 2.3.6 erwähnt.

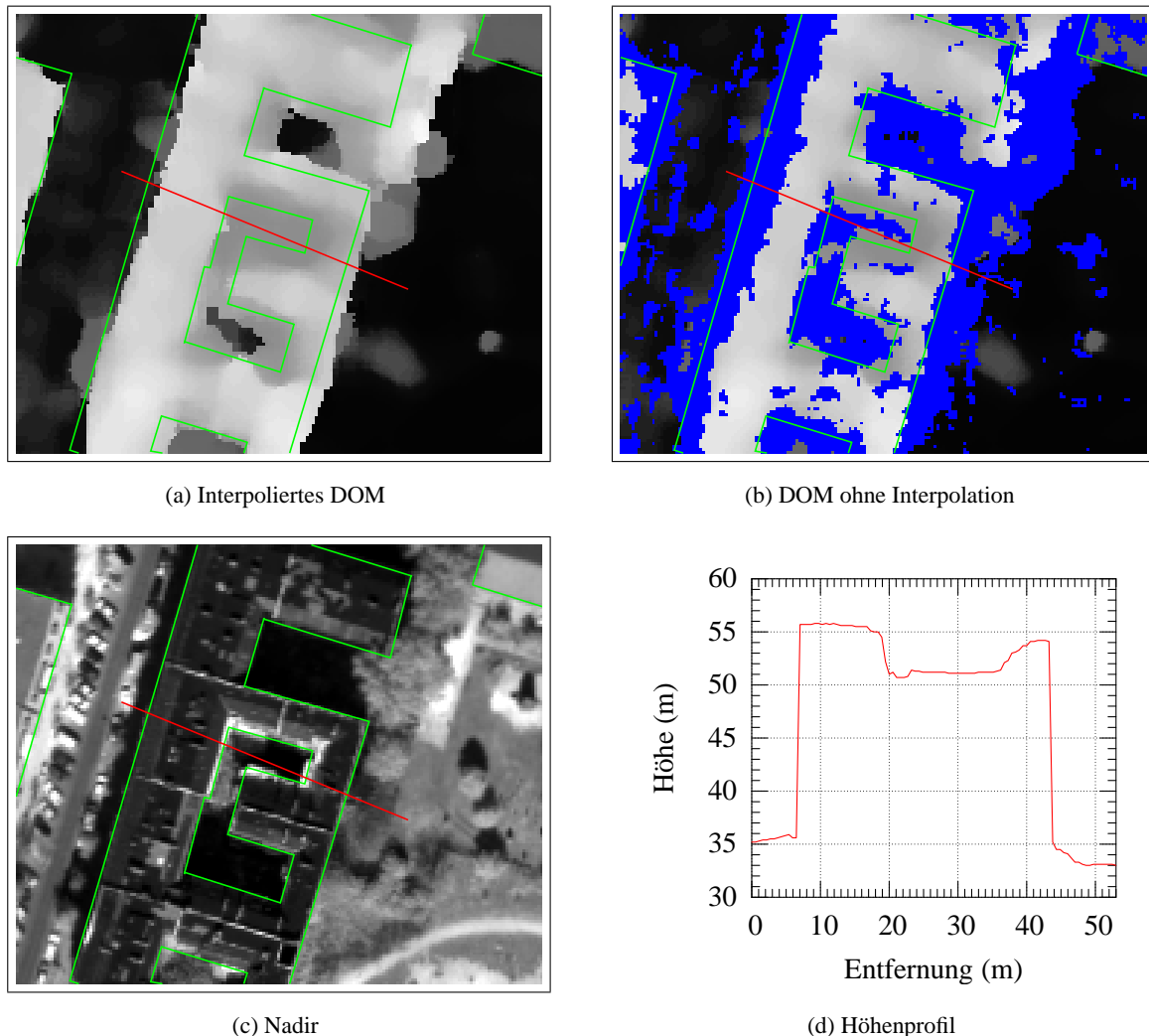


**Abbildung 2.6:** Bildbeispiel zur Interpolation des blickwinkelbedingt nicht erfassten Grundbereichs in der Ortho-Korrektur, zu erkennen an der rot umkreisten Baumgruppe in der Mitte des Bildes.

Bei einer multispektralen Klassifikation wird vorausgesetzt, dass die überlagerten Farbdaten die gleichen Objekte repräsentieren. Bei hochauflösenden Bilddaten, die mit Zeitversatz aufgenommen werden, gilt dies jedoch nicht immer. Sich in der Zeitspanne für die Aufnahme bewegende Objekte werden in einem ortho-korrigierten Farbkomposit i. A. nicht zur Deckung gebracht. Typischerweise ist die Bewegung eigenständig, wie bei Autos oder Bahnen, oder durch Wind bedingt. Hochauflösende Multispektraldaten, die an windigen Tagen aufgenommen wurden, sind beispielsweise bei einer multispektralen Baumklassifikation kritisch zu betrachten. Während sich eine Baumspitze durchaus mehrere Meter im Wind bewegen kann, erfasst ein hochauflösendes Pixel (deutlich) weniger als einen Quadratmeter. In einem solchen Fall werden also unterschiedliche Teile der Szene multispektral überlagert (z. B. ein Blätterpixel mit einem durch einen Ast bestimmten Pixel und ein Bodenpixel). Dies erschwert die Auswertung bzw. kann zu Fehlklassifikationen führen. Ebenso erkennt man in einem HRSC-Farbkomposit sich bewegende Autos durch Mehrfachdarstellung und falsche Farbgebung. Weitere Ausführungen zu dieser Thematik sind bei HOFFMANN (2001, Kapitel 4.6.1) zu finden.

### 2.5.2 Oberflächenmodell

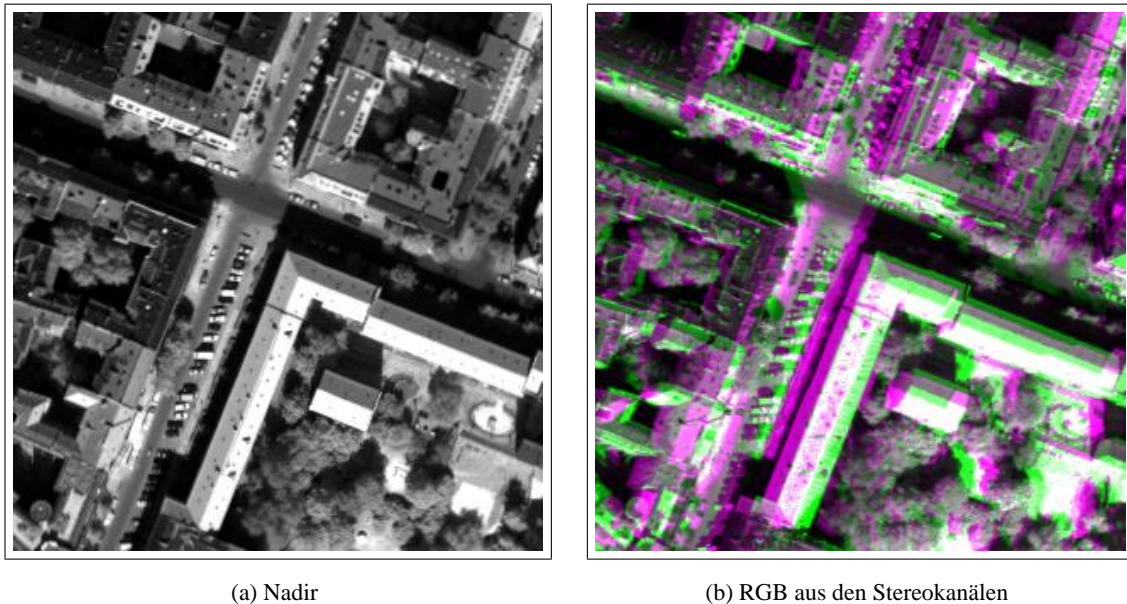
Wie in Abschnitt 2.3.4 beschrieben, gibt es bei der Generierung des Oberflächenmodells mehrere Gründe, weshalb einer DOM-Rasterzelle kein Höhenwert zugeordnet werden kann. An Höhengsprüngen ist dies hauptsächlich durch die Sichtschatten bedingt, wie in Abbildung 2.5 dargestellt. Aus diesem Grund werden zur Berechnung des Höhenmodells fünf Stereokanäle verwendet. Davon müssen mindestens drei einen homologen Punkt definieren. Gegenüber einem Dreizeilenstereoscanner können auf diese Art bei höherem Rechenaufwand mehr Bildpunkte einander zugeordnet werden. Durch



**Abbildung 2.7:** Veranschaulichung des DOM-Interpolationsfehlers. Ausschnitt aus einem HRSC-Oberflächenmodell mit (a) und ohne (b) Interpolation der nicht zugeordneten Punkte (blau überlagert), zugehöriger Nadir-Ausschnitt (c) und ein Höhenprofil (d) entlang der in (a) rot dargestellten Linie. Als Orientierungshilfe sind in diesem Bildausschnitt die Umrisspolygone der Gebäude manuell erfasst worden und in den Abbildungen (a)–(c) grün überlagert.

„Straßenschluchten“ bedingter Sichtschatten wird aber auch hier zu einem Versagen der Bildzuordnung führen. Um ein vollständiges Höhenraster zu erhalten, werden die Höhenwerte der ungefüllten Rasterzellen im DOM anhand ihrer Nachbarn interpoliert. Diese Interpolationsungenauigkeit muss bei einer DOM-basierten Objektextraktion berücksichtigt werden.

In Abbildung 2.7 werden Interpolationsfehler in einem HRSC-DOM exemplarisch dargestellt. Es ist leicht zu sehen (Abb. 2.7(b)), dass vorwiegend an Höhengsprüngen keine Höheninformation ermittelt werden konnte. Die überlagerten Umrisspolygone der Gebäude verdeutlichen diesen Zusam-



**Abbildung 2.8:** Veranschaulichung des Kanalversatzes nach der Ortho-Korrektur. Ausschnitt aus dem Nadirkanal (a) und entsprechende RGB-Darstellung (b) aus den beiden Stereokanälen.

menhang. Außerdem wird der schmale und steil abfallende Innenhof nicht von allen Stereokanälen erfasst. Aufgrund dieses Sichtschattens existieren keine homologen Punkte. Das Ergebnis der Interpolation der nicht-zugeordneten Punkte ist in Abbildung 2.7(a) dargestellt. Das Höhenprofil quer durch den Innenhof (Abb. 2.7(d)) macht deutlich, dass die Höhenwerte hier nicht den tatsächlich vorliegenden Grund widerspiegeln.

Für die Orthokorrektur der Bilddaten hat ein fehlerhaft bestimmter Höhenwert zur Folge, dass ein Bildpunkt nicht auf die korrekte Stelle, sondern leicht verschoben abgebildet wird. Durch die unterschiedlichen Blickwinkel der Farbkanäle ist der Lokalisationsfehler unterschiedlich stark ausgeprägt. Deutlich wird dieser Effekt bei der Überlagerung der Farbkanäle zu einem Komposit. Höhengsprünge weisen dabei geringe Farbsäume auf; die Höhenkanten erscheinen also an leicht unterschiedlichen Positionen in den Farbkanälen. Abbildung 2.8 verdeutlicht dieses Phänomen anhand eines „Extrembeispiels“. Hier sind die ortho-korrigierten Daten der beiden äußeren Kanäle, der Stereosensoren (Blickwinkelunterschied:  $37,8^\circ$ ), in Form eines RGB-Komposits überlagert (Stereo vorne auf R und B, Stereo hinten auf G). Da die beiden Stereosensoren dieselbe spektrale Empfindlichkeit besitzen, treten in der Überlagerung lediglich die Bereiche mangelnder Übereinstimmung farblich hervor (abgesehen von blickwinkelbedingten Effekten). An allen größeren Höhenunterschieden ist ein solcher Versatz zu beobachten. Die Flugrichtung geht bei dieser Aufnahme von Ost nach West, hier also von rechts nach links. In violetten Bereichen ist die Information des vorderen Stereokanals dominant, in grünen Bereichen die des hinteren Stereokanals.

In der Praxis werden üblicherweise nur die Farbkanäle zu einem Komposit überlagert. Deswegen

ist dieses Phänomen dann weniger stark ausgeprägt. Durch die deutlich engeren Blickwinkel der Farbkanäle in der HRSC-AX001 werden diese Effekte abermals reduziert.

Aufgrund der genannten Probleme mit Sichtschatten bei Höhengsprüngen ist eine einfache und exakte Extraktion erhöhter Objekte aus dem DOM nicht zu erwarten. Der naheliegende Ansatz (vorausgesetzt das Terrain ist zu vernachlässigen) durch Höenschwellen die erhöhten Objekte zu identifizieren, wird nur bei perfekten Höhenmodellen zufriedenstellende Ergebnisse liefern. Durch einen Höenschwellwert detektierte Objektkonturen sind besonders deutlich von dem durch die Interpolation bedingten Lokalisationsfehler betroffen.

Die geschilderten Unzulänglichkeiten im DOM sind eine zentrale Motivation für das in dieser Arbeit geschilderte Vorgehen. Im folgenden Kapitel wird der Ansatz, aus einem Oberflächenmodell unter Zuhilfenahme der Multispektralkanäle Objekte zu extrahieren, vorgestellt. Insbesondere die Objektkontur soll dabei trotz möglicher Ungenauigkeiten im DOM genau lokalisiert werden.





## 3 Extraktion erhöhter Objekte aus einem Oberflächenmodell

Viele der in einer Interpretation von Fernerkundungsdaten relevanten Objekte heben sich durch ihre Höhe von der Umwelt ab. Insbesondere gilt dies für die im Mittelpunkt der Arbeit stehenden Gebäude. Die Höheninformation stammt in dieser Arbeit aus digitalen Oberflächenmodellen (DOMs). Wie in Abschnitt 1.2 ausgeführt, vereinigt ein Oberflächenmodell Grund- und Objekthöhen. Ziel dieses Kapitels ist die Trennung dieser zwei Arten von Höheninformation, also die Berechnung eines Terrainmodells (DTM) und eines auf Grundhöhe normierten Oberflächenmodells (nDOM) aus dem DOM.

Einführend wird in Abschnitt 3.1 ein Überblick über bestehende Ansätze zur DTM-Berechnung aus einem DOM gegeben und das eigene Vorgehen motiviert. Zur Berechnung des DTM werden in den Abschnitten 3.2 und 3.3 zwei Verfahren vorgestellt. In Abschnitt 3.4 werden die beiden Herangehensweisen einander gegenüber gestellt und die damit erzielten Ergebnisse bewertet.

Nach einer Normierung des Oberflächenmodells können die erhöhten Regionen extrahiert werden, was in Abschnitt 3.5 behandelt wird. Resultat der in diesem Kapitel vorgestellten Methoden sind Segmente erhöhter Regionen mit bestimmten minimalen Abmessungen.

### 3.1 Einführung und Literaturüberblick

Grundlegende Annahme aller Verfahren zur DOM-Normierung ist zum einen, dass im DOM auch Grundpunkte vorliegen, die dann zur Bestimmung der Grundhöhe bei erhöhten Objekten referenziert werden. Bei fensterbasierten Verfahren stellt sich die Frage nach der Wahl der Fenstergröße. Das Fenster muss groß genug genug sein, um zu gewährleisten, dass immer ein Grundpunkt in der Nachbarschaft enthalten ist. Also muss das Fenster größer als das größte im DOM vorkommende Objekt sein. Andererseits sollte das Fenster möglichst klein sein, um die Verwendung objektferner Punkte bei der Bestimmung der Grundhöhe zu vermeiden. VOSSELMAN (2000) bestimmt die Fenstergröße über eine maximal vorgegebene Steigung oder aber über einen Trainingsdatensatz.

Eine weitere, allen Herangehensweisen gemeinsame Annahme ist, dass die lokalen Höhengsprünge, also große Höhenänderungen auf kleiner Distanz, durch die Objekte über Grund hervorgerufen werden. Von der Morphologie des Geländes hingegen wird angenommen, dass sie schwächeren Änderungen unterliegt. Bei einer Abweichung von dieser Annahme, z. B. bei sehr zerklüftetem Terrain,

werden alle Verfahren zur DOM-Normierung unbefriedigende Ergebnisse liefern. Jedoch gibt es Ansätze, geomorphologisch relevante Strukturen wie Bruchkanten in das aus einer Normierung abgeleitete DTM zu modellieren (KRAUS UND PFEIFER 2001). Voraussetzung dafür ist jedoch, dass die Position dieser Strukturen in etwa bekannt ist.

Ein verbreitetes Vorgehen zur Normierung von Oberflächenmodellen ist der Einsatz von Grauwert-Morphologie (WEIDNER UND FÖRSTNER 1995, VOSSELMAN 2000), bzw. allgemeiner, *Rangordnungsoperatoren*. Das Vorgehen beim morphologischen Öffnen (auch Minimum/Maximum-Rangordnungsoperator) ist wie folgt (s. HARALICK ET AL. 1987): Zuerst wird jedem Bildpunkt mit einem Minimumfilter das Minimum in einer definierten Nachbarschaft zugewiesen. Dies entspricht einer morphologischen Erosion. Anschließend wird jedem Bildpunkt des Resultats das Maximum in der Nachbarschaft zugewiesen, was einer morphologischen Dilatation entspricht. Mit dem ersten Schritt sollen die erhöhten Objekte entfernt und im zweiten Schritt die Grundhöhe wieder angepasst werden.

Da die Minimum- und Maximumoperatoren anfällig gegenüber Fehlern (Ausreißern) im DOM sind, schlagen ECKSTEIN UND MUNKELT (1995) eine Modifikation der zweifachen Rangordnungsoperatoren vor: Befinden sich in der Nachbarschaft<sup>1</sup>  $n$  Bildpunkte, so wird im ersten Durchgang nicht das erste Element der geordneten Höhenwerte gewählt (Erosion), sondern das  $k$ -te. Invers dazu wird im zweiten Durchgang nicht das letzte Element (Dilatation), sondern das  $n - k$ -te gewählt.

Verfahren zur DTM-Bestimmung mittels Rangordnungsoperatoren berechnen für *jeden Bildpunkt* eine Terrainhöhe. Somit wird auch real vorhandenen Grundpunkten eine neue Höhe zugeordnet. Die im DOM vorhandene Höhe realer Grundpunkte wird bei Vorliegen einer Geländesteigung sogar verzerrt.

SCHAEFFER (1999) benutzt eine doppelte Rangordnungsoperation nur dazu, ein grobes DTM abzuleiten. Aus den Höhenunterschieden zum DOM ergeben sich die erhöhten Objekte. Für sie wird die Grundhöhe anhand benachbarter Grundpunkte linear interpoliert. Dieser zweite Teil entspricht dem zweiten Teil des in Abschnitt 3.2 vorgestellten Verfahrens zur Ermittlung der Grundhöhe.

Eine anderer Ansatz zur Ermittlung des DTM ist die *lineare Vorhersage* (KOCH 1999, LOHMANN UND KOCH 1999, KRAUS UND PFEIFER 2001). Zuerst wird in die DOM-Ausgangsdaten eine Grundebeine bzw. ein Polynom niedriger Ordnung eingepasst. Aus dem Unterschied der ursprünglichen DOM-Höhen zu der so eingepassten Grundebeine wird für jeden Punkt eine Gewichtung abgeleitet. Geringe Höhendifferenzen führen zu einem hohen Gewicht und umgekehrt. Entsprechend dieser Gewichtung wird die Grundebeine iterativ immer wieder neu eingepasst und die Gewichte neu bestimmt. Zudem werden Punkte mit zu großem Abstand zur Ausgleichsebene als „erhöht“ eingestuft und bei künftigen Ausgleichsiterationen ausgeschlossen. KRAUS UND PFEIFER (2001) wiederholen diesen Prozess auf mehreren Auflösungsebenen, so dass Filterung und Interpolation integriert erfolgen.

---

<sup>1</sup> Korrekter Terminus bei morphologischen Operatoren: „strukturierendes Element“

Prinzipiell gilt wie bei den Rangordnungsoperatoren auch für die lineare Vorhersage, dass alle Bildpunkte behandelt werden. Durch das iterative Einpassen und Korrigieren wird der Approximationsfehler jedoch schrittweise verringert.

Weitere in der Literatur erwähnte Methoden zur DOM-Normierung basieren auf der konvexen Hülle des DOM, der Approximation der DOM-Oberfläche durch eine Spline-Fläche oder Filter zur Glättung des DOM.

In den beiden nachfolgenden Abschnitten werden zwei Herangehensweisen zur Ableitung der Terrainhöhe aus einem Oberflächenmodell herausgearbeitet. Der erste Ansatz, bezeichnet als „Interpolation der Grundhöhe“ (MAYER 2000a,b), ist in der Detektion der erhöhten Punkte der Grauwert-Morphologie ähnlich. Jedoch wird *nur der Höhenwert der als erhöht eingestuften Punkte verändert*; die Grundpunkte bleiben unverändert. Zur Wahl der Fenstergröße für die Referenzierung mit benachbarten Grundpunkten wird die *Kombination mehrerer Parameterkonstellationen* vorgeschlagen.

Motivation für ein weiteres Verfahren ist das Bestreben, einen höheren Automatisierungsgrad zu erreichen. Unter Veränderung der DOM- bzw. DTM-Auflösung wird bei der „pyramidalen Verfeinerung der Geländehöhe“ ein Geländemodell automatisch berechnet. Die Detektion erhöhter Punkte basiert dabei auf der Vorgabe einer maximalen Steigung des Geländes und einer maximalen Objekt-abmessung. Aus diesen Werten werden in jeder Auflösungsebene die passenden Fenstergrößen bzw. Schwellwerte berechnet.

## 3.2 Interpolation der Grundhöhe

Als Referenz für die Interpolation der Grundhöhe bei erhöhten Objekten dienen möglichst nahegelegene Grundpunkte. Die Unterscheidung von Grundpunkt und erhöhtem Punkt wird dadurch getroffen, dass ein bestimmter Höhenunterschied vorliegt. Dazu wird das aktuell betrachtete Pixel dem minimalen Höhenwert in einer Nachbarschaft  $N^{(gi)}$  gegenübergestellt. Die Größe der Nachbarschaft  $r_{gi}$  (Suchradius für Minimumabgleich) ist dabei ein Eingabeparameter. Einerseits sollte sie möglichst klein sein, um zu gewährleisten, dass die Unterscheidung in Grundpunkte und erhöhte Punkte anhand möglichst naher Punktpaare erfolgt. Andererseits muss sie so groß sein, dass auch von dem ausgedehntesten erhöhten Objekt noch ein Grundreferenzpunkt erreicht werden kann. Der Nachbarschaftsradius ist in Abhängigkeit von der Bildauflösung als Standardwert vorgegeben, kann vom Benutzer aber frei variiert werden. So kann auch auf die Besonderheiten der Szene eingegangen werden. Eine geschlossene Walddecke erfordert beispielsweise einen höheren Suchradius als eine aus vielen freistehenden Häusern bestehende Neubausiedlung.

Für jedes DOM-Pixel wird nun innerhalb von  $N^{(gi)}$  das Höhenminimum bestimmt und zur Nor-

mierung des aktuellen Pixels herangezogen<sup>2</sup> (Algorithmus 1, Zeile 2):

$$\text{nDOM}'(i) = \text{DOM}(i) - \min_{k \in N_i^{(\text{gi})}} \text{DOM}(k) \quad (3.1)$$

Resultat ist ein vorläufiges normiertes Oberflächenmodell  $\text{nDOM}'$ . Der dafür notwendige zweidimensionale Suchvorgang ist separierbar in zwei eindimensionale Suchschritte. Dadurch reduziert sich der Rechenaufwand von  $O(n^2)$  auf  $O(n \log n)$  mit  $n$  proportional zum Suchradius. Da die Suchfenster in der Praxis relativ groß sind (Durchmesser/Kantenlänge von 100–200 Pixel), ist die Beschleunigung des Programmablaufs durch die Separierung erheblich.

**Eingabe:** DOM

**Ausgabe:** DTM

DOM2DTM\_GI(DOM)

- (1) DTM  $\leftarrow$  DOM
- (2)  $\text{nDOM}' \leftarrow \text{RELATIVEHEIGHT}(\text{DOM}, r_{\text{gi}})$
- (3)  $\mathcal{S}_{\text{gi}} \leftarrow \text{SEGMENTHIGH}(\text{nDOM}', \theta_{\Delta h})$
- (4)  $\mathcal{S}_{\text{gi}} \leftarrow \text{DILATE}(\mathcal{S}_{\text{gi}})$
- (5) **foreach**  $\text{seg}_i \in \mathcal{S}_{\text{gi}}$
- (6)      $\mathcal{T} \leftarrow \text{TRIANGULATE}(\text{seg}_i)$
- (7)     DTM( $\text{seg}_i$ )  $\leftarrow \text{INTERPOLATE}(\text{DOM}(\text{seg}_i), \mathcal{T})$
- (8) **return** DTM

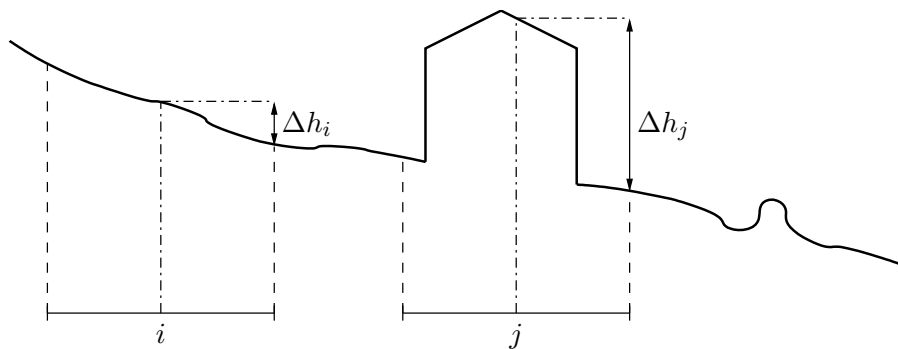
**Algorithmus 1:** Berechnung eines DTM durch Grundhöheninterpolation.

Ist die Höhendifferenz von aktuellem Pixel und Minimum in dessen Nachbarschaft größer als ein vorgegebener Höschwellwert  $\theta_{\Delta h}$ , so wird das aktuelle Pixel als „erhöht“ markiert (in Algorithmus 1 als Segmentierung  $\mathcal{S}_{\text{gi}}$  bezeichnet), andernfalls als Grundpixel. Dies ist in Abbildung 3.1 anhand eines eindimensionalen Bildbeispiels dargestellt. An den Stellen  $i$  und  $j$  sind jeweils die Abmessungen des Nachbarschaftsfensters aufgetragen. Die durch das Minimum in diesem Fenster gegebene Referenzhöhe führt zu der Höhendifferenz  $\Delta h_i$  bzw.  $\Delta h_j$ . Dabei liegt  $\Delta h_j$  über der vorgegebenen Höschwelle, womit der Punkt  $j$  als „erhöht“ eingestuft wird.

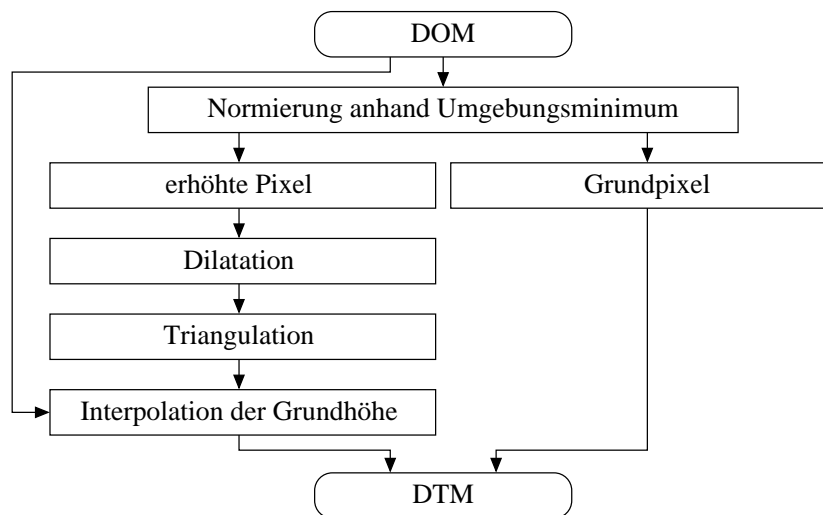
In diesem Ansatz wird die Höhe von Grundpunkten im Unterschied zur DTM-Bestimmung mittels Grauwert-Morphologie nicht verändert. Lediglich die Grundhöhen der ermittelten erhöhten Regionen werden nun anhand benachbarter Grundpixel interpoliert. Dazu sollten verlässliche Grundpixel herangezogen werden, also möglichst nicht von Interpolationsfehlern des Oberflächenmodells betroffene Punkte (s. Abschnitt 2.3.4 und 2.5.2). Aus diesem Grund werden die Regionen erhöhter Pixel durch eine morphologische Operation dilatiert, bzw. die Grundregionen entsprechend erodiert (Algorithmus 1, Zeile 4).

---

<sup>2</sup>In Gleichung 3.1 sind die Höhenmodelle zugunsten der Übersichtlichkeit nur in einer Dimension indiziert.



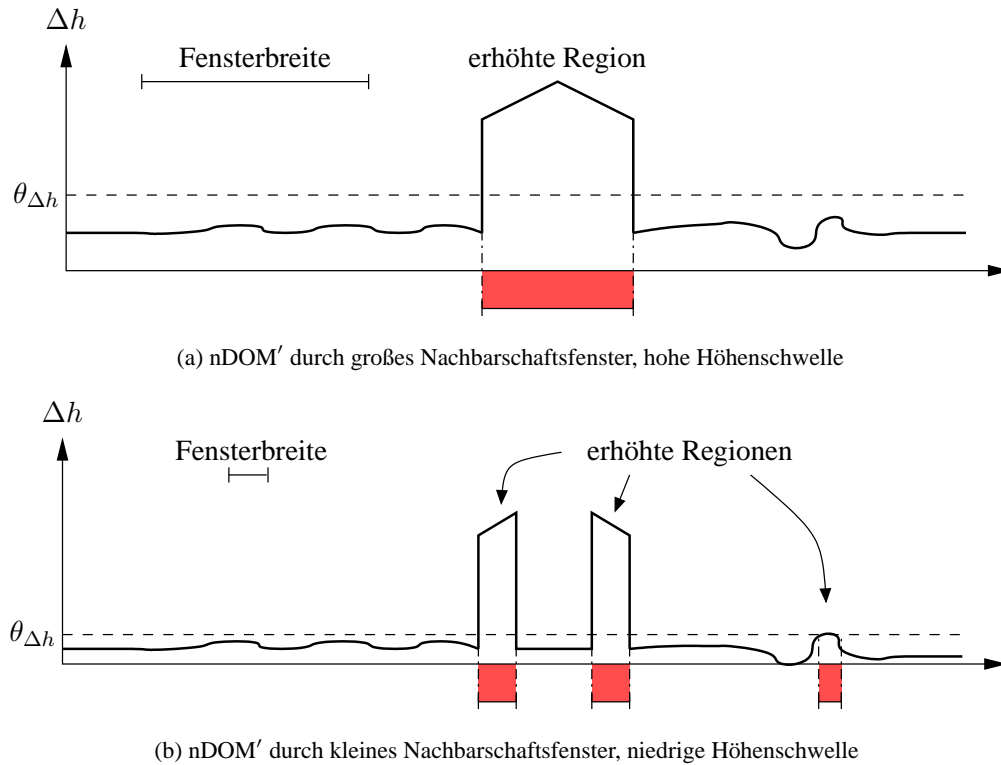
**Abbildung 3.1:** Exemplarische Veranschaulichung des Vorgehens zur Detektion erhöhter Punkte für die Grundhöheninterpolation an einem Höhenprofil.



**Abbildung 3.2:** Schematische Darstellung des Ablaufs der Grundhöheninterpolation anhand benachbarter Grundpunkte.

Um die Grundinterpolation der erhöhten Regionen möglichst effizient zu gestalten, werden deren Umrisspolygone nach dem Delaunay-Algorithmus trianguliert (s. z. B. LEE UND SCHACHTER 1980). Die Kreisfläche um jedes bei der Triangulation entstehende Dreieck enthält bei Delaunay-triangulierten Flächen keine weiteren Eckpunkte. Dadurch ist gewährleistet, dass die interpolierten Werte nur aus nahen Eckpunkten berechnet werden.

Durch die vorhergehende Dilatation der erhöhten Regionen ist jeder Triangulationspunkt ein „verlässlicher“ Grundpunkt. In Algorithmus 1 ist die Menge der Triangulationspunkte, die ein erhöhtes Segment  $\text{seg}_i \in \mathcal{S}_{\text{gi}}$  beschreiben, mit  $\mathcal{T}$  bezeichnet. Die triangulierten Flächen werden nun anhand der drei jeweiligen Triangulationsgrundpunkte linear interpoliert. Kombiniert man die interpolierten Grundflächen mit den vorab segmentierten Grundpunkten, so erhält man ein vollständig gefülltes Grundhöhenraster, das digitale Terrainmodell (DTM). Die segmentweise Triangulation und Interpo-



**Abbildung 3.3:** Auswirkung der Größe des Nachbarschaftsfensters und der Höhengschwelle auf die Detektion erhöhter Regionen. Grundlage dafür ist das in Abb. 3.1 dargestellte Höhenprofil.

lation ist in Algorithmus 1 in den Zeilen 5–7 wiedergegeben. Das gesamte Vorgehen ist in Abbildung 3.2 schematisch dargestellt.

Die wesentlichen Parameter sind bei diesem Verfahren die Größe der Suchmaske für den Minimumabgleich  $r_{gi}$  sowie der Höhengschwellwert für die Segmentierung erhöht/Grund  $\theta_{\Delta h}$ . Die Festlegung auf ein Parameterpaar ist oft zu einschränkend. Einerseits führt die Wahl einer relativ hohen Höhengschwelle dazu, dass alle erhöhten Objekte unterhalb dieses Schwellwerts als Grund behandelt werden und somit im resultierenden DTM in ihrer vollen Höhe erhalten bleiben. Eine niedrige Höhengschwelle hingegen muss an einen kleinen Suchradius gekoppelt sein, um auszuschließen, dass fast alle Pixel als erhöht betrachtet werden (aufgrund geringer Terrainschwankungen). Dann jedoch kann das Innere von ausgedehnten erhöhten Objekten nicht mehr anhand eines nahen Grundpunkts normiert werden. Ausgedehnte erhöhte Objekte können somit nicht mehr als solche identifiziert werden. Dieser Zusammenhang von Fenstergröße und Höhengschwelle ist in Abbildung 3.3 an einem Höhenprofil dargestellt, das aufgrund des Höhenminimums in einem Nachbarschaftsfenster vorläufig normiert wurde. Zur Detektion ausgedehnter erhöhter Objekte ist ein großes Nachbarschaftsfenster und somit auch ein relativ hoher Höhengschwellwert nötig (Abbildung 3.3(a)). Niedrigere erhöhte Objekte können mit einem kleinen Nachbarschaftsfenster und einer geringeren Höhengschwelle de-

tektiert werden (Abbildung 3.3(b)).

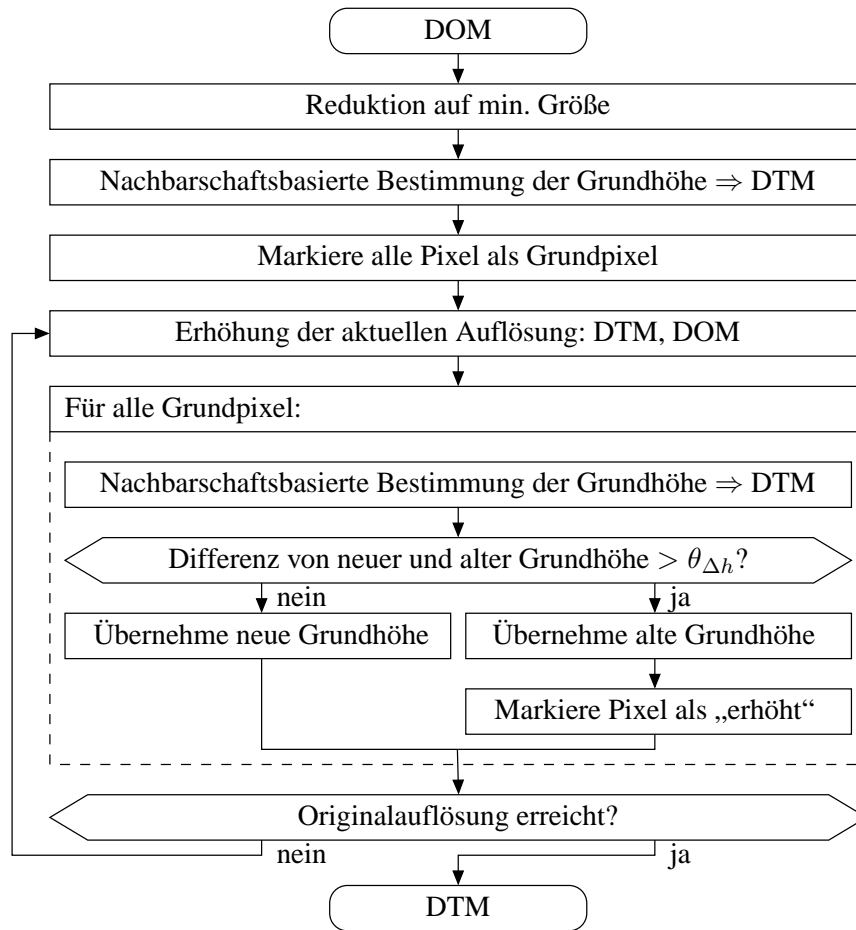
Schlussfolgerung daraus ist es, möglichst eine Kombination dieser beiden Parameterkonstellationen anzuwenden. Einmal eine Grundpixelsegmentierung anhand einer hohen Höschwelle und großem Suchfenster zur Erfassung großer Höhenunterschiede; ein weiteres Mal eine Segmentierung mit kleiner Höschwelle und kleinem Suchfenster zur Erfassung niedrigerer und kleinerer erhöhter Objekte. Kombiniert werden die beiden Segmentierungsergebnisse durch eine logische UND-Verknüpfung. Das weitere Vorgehen ist unverändert wie bereits geschildert.

Die Segmentierung erhöhter Regionen  $S_{gi}$  wird bei diesem Verfahren als Zwischenprodukt abgelegt. So ist es dem Benutzer im Bedarfsfall möglich, durch Manipulation des Segmentierungsergebnisses korrigierend einzugreifen. Auch externes Wissen über die Geländebeschaffenheit kann so einfließen.

### 3.3 Pyramidale Verfeinerung der Geländehöhe

Um bei dem letztgenannten Verfahren erfolgreich sowohl kleine wie auch große Höhsprünge detektieren zu können, wird die Kombination der Ergebnisse aus unterschiedlichen Parameterkonstellationen vorgeschlagen. Bei dem nun zu behandelnden Algorithmus steht eine automatisierte Parameteradaption im Vordergrund. Die eingangs erwähnte Vereinigung von Terrain- und Objekthöhe im DOM wird mittels Vorgabe der maximalen Geländesteigung aufgebrochen. Tritt eine Steigung auf, die größer als die Vorgabe ist, so wird sie einem erhöhten Objekt und nicht dem Terrain zugeschrieben. Wie im letztgenannten Verfahren zur Grundinterpolation wird die Grundhöhe durch den Vergleich zweier Höhenwerte innerhalb eines Fensters bestimmt. Die realen Abmessungen des Suchfensters werden jedoch durch Skalierung des Höhenmodells in einem pyramidalen Ansatz variiert. Einen Überblick über das Vorgehen geben Abbildung 3.4 und Algorithmus 2. Die einzelnen Schritte werden nun diskutiert.

Die Bestimmung der Grundhöhe erfolgt pixelweise mittels eines Nachbarschaftsfensters (Algorithmus 2, Zeilen 4 und 10). Alle Höhenwerten in diesem Fenster werden in einem Histogramm aufgetragen. Nach einer Glättung des Histogramms (durch Mittelwertfilterung in einer Umgebung von fünf Einträgen) dient die Position des ersten Maximums als Grundhöhe für das aktuell betrachtete Pixel. Abbildung 3.5 verdeutlicht das Vorgehen an einem Bildbeispiel. Zu den drei exemplarisch im DOM markierten Ausschnitten werden die jeweiligen Höhenhistogramme gezeigt (rot: Histogramm, blau: Glättung). Die ersten Maxima, die als lokal bestimmte Grundhöhe herangezogen werden, sind cyanfarben markiert. Im ersten Fall werden nur Dachpixel herangezogen. Bereits der erste Eintrag stellt ein (schwach ausgeprägtes) lokales Maximum dar. Mit dieser Fenstergröße kann also kein Grundpunkt referenziert werden. Im zweiten Fall sind nur Grundpixel vorhanden, die Grundhöhenbestimmung kann also nicht fehlschlagen. Der dritte Fall zeigt, dass bereits mit wenigen Grundpixeln in der jeweiligen Nachbarschaft die Grundhöhe eines erhöhten Objekts erfolgreich bestimmt werden



**Abbildung 3.4:** Schematische Darstellung des Ablaufs der pyramidalen Verfeinerung der Grundhöhe.

kann.

Im Unterschied zu Rangordnungsoperatoren wird die Grundhöhe somit nicht anhand eines fixen Rangs der geordneten Höhenwerte bestimmt, sondern variabel in Abhängigkeit vom ersten Maximum. Dadurch wird eine größere Robustheit gegenüber vereinzelt Höhenminima erreicht als mit dem im letzten Verfahren zum Einsatz kommenden Minimumabgleich<sup>3</sup>.

Das Oberflächenmodell wird zur Berechnung der Grundhöhe jedoch nicht nur einmal herangezogen, sondern mehrmals in Form einer Auflösungshierarchie. Aus der festzusetzenden Anzahl an Iterationen  $K$ , dem DOM-Skalierungsfaktor  $s_d$  und der DOM-Originalpixelgröße  $x_d$  ergeben sich

---

<sup>3</sup>Die beiden Methoden zur Ermittlung der Grundhöhe sind nicht charakteristisch für eines der beiden Verfahren. Sowohl Minimumabgleich wie auch histogrammbasierte Ermittlung der Grundhöhe können ebenfalls im jeweils anderen Ansatz zur DTM-Berechnung zum Einsatz kommen.



die Auflösungsstufen des DOM als

$$x_k = x_d \cdot s_d^{K-k}, \quad k = 1, \dots, K. \quad (3.2)$$

**Eingabe:** DOM

**Ausgabe:** DTM

DOM2DTM\_PV(DOM)

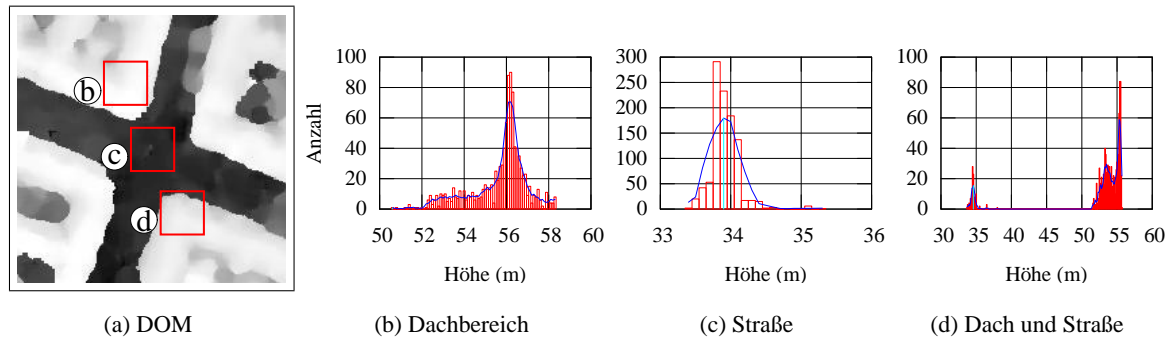
- (1)  $\text{DOM}^{(1)} \leftarrow \text{SUBSAMPLE}(\text{DOM}, s_d^{K-1})$
- (2)  $\mathcal{S}_{\text{pv}}^{(1)} \leftarrow \text{Menge aller Pixel von } \text{DOM}^{(1)}$
- (3) **foreach**  $i \in \mathcal{S}_{\text{pv}}^{(1)}$
- (4)  $\text{DTM}^{(1)}(i) \leftarrow \text{GROUNDHEIGHT}(\text{DOM}^{(1)}, i, r_{\text{pv}} \cdot s_w)$
- (5) **for**  $k \leftarrow 2$  **to**  $K$
- (6)  $\text{DOM}^{(k)} \leftarrow \text{SUBSAMPLE}(\text{DOM}, s_d^{K-k})$
- (7)  $\text{DTM}^{(k)} \leftarrow \text{SUPERSAMPLE}(\text{DTM}^{(k-1)}, s_d)$
- (8)  $\mathcal{S}_{\text{pv}}^{(k)} \leftarrow \text{SUPERSAMPLE}(\mathcal{S}_{\text{pv}}^{(k-1)}, s_d)$
- (9) **foreach**  $i \in \mathcal{S}_{\text{pv}}^{(k)}$
- (10)  $h \leftarrow \text{GROUNDHEIGHT}(\text{DOM}^{(k)}, i, r_{\text{pv}} \cdot s_w)$
- (11) **if**  $|\text{DTM}^{(k)}(i) - h| > \theta_{\Delta h}^{(k)}$
- (12)  $\mathcal{S}_{\text{pv}}^{(k)} \leftarrow \mathcal{S}_{\text{pv}}^{(k)} \setminus \{i\}$
- (13) **else**
- (14)  $\text{DTM}^{(k)}(i) \leftarrow h$
- (15) **return** DTM

**Algorithmus 2:** Berechnung eines DTM durch pyramidale Verfeinerung der Geländehöhe.

Zur Skalierung wird das DOM entsprechend dem Skalierungsfaktor geglättet und unterabgetastet. In Algorithmus 2 (Zeilen 1 und 6) sind die entsprechenden DOM-Auflösungsstufen mit  $\text{DOM}^{(k)}$  bezeichnet, das beschriebene Unter- und das äquivalente Überabtasten mit SUBSAMPLE bzw. SUPERSAMPLE. Aus der Variation der DOM-Auflösung und fixer Größe des Nachbarschaftsfensters folgt somit von Iteration zu Iteration eine Verkleinerung des realen Suchbereichs um den Faktor  $s_d$ . Jedoch besteht auch die Möglichkeit, die Fenstergröße zusätzlich mittels eines Skalierungsfaktors  $s_w$  zu verändern.

Aus der Vorgabe einer maximalen Entfernung für die Normierung durch benachbarte Grundpunkte,  $d_{\text{max}}^{\text{ground}}$ , und der maximal erlaubten Steigung des Geländes,  $m_{\text{max}}^{\text{ground}}$ , ergibt sich eine Höhenschwellwert für Grundpunkte. Durch Skalierung mit der Auflösung in der jeweiligen Pyramidenstufe resultieren die Höhenschwellen  $\theta_{\Delta h}^{(k)}$  mit

$$\theta_{\Delta h}^{(k)} = \frac{d_{\text{max}}^{\text{ground}} m_{\text{max}}^{\text{ground}}}{s_d^{K-k} s_w^{K-k}}, \quad k = 1, \dots, K. \quad (3.3)$$

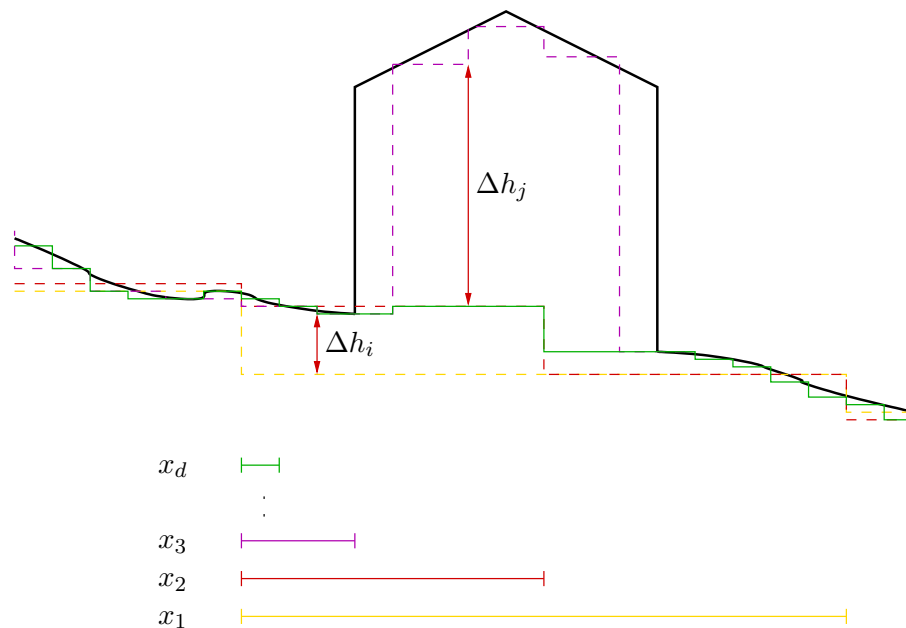


**Abbildung 3.5:** Fallbeispiele zur histogrammbasierten Bestimmung der Grundhöhe. Aus den im DOM (a) rot eingezeichneten Ausschnitten werden die Höhenhistogramme (c–d) berechnet.

Liegt die Höhendifferenz von aktuell betrachtetem Pixel zu der aus dem Nachbarschaftshistogramm ermittelten Grundhöhe unter dieser Schwelle, so gilt das aktuelle Pixel als Grundpixel (in Algorithmus 2 über die Zugehörigkeit zur Segmentierung  $S_{pv}$  implementiert). Die Höhe der Grundpixel wird in den nachfolgenden Iterationen durch Erhöhung der DOM-Auflösung immer genauer bestimmt. Andernfalls wird das aktuelle Pixel als „erhöht“ gekennzeichnet (in Algorithmus 2, Zeile 12). Voraussetzung für eine erfolgreiche DTM-Berechnung ist die korrekte Wahl der Fenstergröße: Für die initiale DOM-Auflösung  $x_1$  muss das Fenster so groß sein, dass damit immer Grundpixel abgedeckt werden. Dadurch ist gewährleistet, dass im ersten Schritt ein zwar grobes, aber in dieser Auflösungsstufe korrektes DTM erzeugt werden kann. Dieses initiale DTM ist Grundlage für die nachfolgenden Iterationsschritte, in denen es immer weiter verfeinert wird. Genauer gesagt, erfolgt die Verfeinerung nur für die Pixel, die noch als Grundpixel gelten. Wird ein Pixel einmal als „erhöht“ eingestuft, so behält es seine im vorhergehenden Iterationsschritt bestimmte Grundhöhe bei. Veranlassung dafür ist, dass die reale Größe der Nachbarschaft in jeder nachfolgenden Iteration kleiner wird. Da der Höhenschwellwert aber bereits in der aktuellen Iteration überschritten wird (d. h. im Höhenhistogramm gibt es keine wesentlich kleineren Höhenwerte), werden bei einer Verkleinerung der Nachbarschaft auch keine Grundpixel mehr abgedeckt (d. h. die Nachbapixel befinden sich alle *auf* einem erhöhten Objekt).

Der Verfeinerungsprozess wird solange iteriert, bis die Originalauflösung des DOM erreicht ist (Algorithmus 2, Zeilen 5–14). In der Praxis werden 3–8 Iterationen zu zufriedenstellenden Ergebnissen führen. Resultat ist ein DTM in gleicher Auflösung wie das ursprüngliche DOM.

Zur Veranschaulichung des Vorgehens ist in Abbildung 3.6 ein eindimensionales Bildbeispiel dargestellt. Die farbig dargestellten Linien geben die Kantenlänge bzw. den Durchmesser des Nachbarschaftsfensters der jeweiligen Auflösungsstufe an. Das initiale DTM ergibt sich aus der gelb gestrichelt eingezeichneten Linie. In den nachfolgenden Iterationen wird dieses grobe DTM immer weiter verfeinert. Beim Übergang von der Ebene mit Pixelauflösung  $x_2$  zu der mit Auflösung  $x_3$  sind exemplarisch an den Positionen  $i$  und  $j$  die Höhendifferenzen der jeweiligen Grundhöhen durch Pfeile



**Abbildung 3.6:** Exemplarische Veranschaulichung des Vorgehens zur pyramidalen Verfeinerung der Grundhöhe an einem Höhenprofil. Die farbig dargestellten Linien geben die Kantenlänge des Nachbarschaftsfensters und die daraus resultierende Grundhöhe in der jeweiligen Auflösungsstufe an.

symbolisiert. An Stelle  $i$  ist die Differenz kleiner als der Schwellwert, deshalb wird Grundhöhe noch in weiteren Iterationen verfeinert. An Stelle  $j$  hingegen werden in der dritten Auflösungsstufe nur noch Referenzpunkte auf dem Gebäudedach erreicht. Da in der zweiten Auflösungsstufe auch echte Grundpunkte im Nachbarschaftsfenster enthalten sind, entsteht an Stelle  $j$  eine wesentlich höhere Höhendifferenz. Deshalb werden die Grundhöhen hier nicht weiter verfeinert, sondern verbleiben auf dem Niveau der zweiten Auflösungsebene.

## 3.4 Evaluation der DTM-Berechnung

Nach einer allgemeinen Gegenüberstellung der beiden vorgestellten Ansätze zur Berechnung eines DTM aus einem DOM (Abschnitt 3.4.1) werden diese anhand eines Referenzdatensatzes qualitativ und quantitativ beurteilt (Abschnitt 3.4.2).

### 3.4.1 Vergleich der beiden Ansätze

Das Verfahren zur Interpolation der Grundhöhe ist prinzipiell in dem gleichen Maße automatisierbar wie das der pyramidalen Verfeinerung. Jedoch steht hier im Vordergrund, über die bei der Verarbeitung anfallenden Zwischenprodukte dem Nutzer die Möglichkeit zur Interaktion zu geben. So können bereits existierende Segmentierungen in erhöhte Regionen und Grund integriert werden. Ferner kann

Kriterium	Grundinterpolation	Pyramidale Verfeinerung
Automatisierungsgrad	mittel	hoch
Manuelle Interaktion	gut möglich	nicht angedacht
Verarbeitungszeit	mittel	mittel – sehr hoch (je nach Anzahl an Iterationen)
Genauigkeit (bei geeigneter Parameterwahl)	gut, durch manuelle Interaktion können grobe Fehler ausgeschlossen werden	gut – mittel

**Tabelle 3.1:** Gegenüberstellung der Vor- und Nachteile der beiden Verfahren zur DOM-Normierung.

durch den Nutzer die Qualität der Segmentierung durch Überlagerung mit Bilddaten visuell relativ einfach abgeschätzt werden. Fehlzugeordnungen können so noch vor der eigentlichen Interpolation erkannt und behoben werden. Zudem besteht die Möglichkeit durch Spezifikation entsprechender Umrisspolygone bestimmte Regionen explizit als Grundpunkte auszuschließen. Das ist einerseits dann sinnvoll, wenn vorab bekannte Fehler im DOM nicht auch im DTM auftauchen sollen. Andererseits kann es erwünscht sein, nicht alle Vertiefungen im DOM als Grund, also mit Objekthöhe Null, zu behandeln. Temporären Vertiefungen wie Baugruben im auf Grundhöhe normierten DOM (nDOM) eine negative Höhe zuzuweisen, erscheint sinnvoll. Dies wird aber erst möglich, wenn solche Regionen bei der Grundpunktfindung ausgeschlossen sind.

Gemeinsam ist beiden Ansätzen, dass die Entscheidung, ob für das aktuell betrachtete Pixel ein Grundpunkt vorliegt, für jede Parameterkonstellation neu getroffen wird. In der pyramidalen Verfeinerung wird jedoch auf die Ergebnisse der vorangegangenen Stufe aufgebaut, während in der Grundinterpolation die Entscheidungen unabhängig voneinander getroffen werden. Hier ist der Nutzer verantwortlich für eine sinnvolle Parameterkonstellation.

Anstelle einer direkten Spezifikation der maximalen Geländesteigung bei der pyramidalen Verfeinerung, ergibt sich bei der Grundinterpolation die Steigung nur indirekt aus dem Zusammenspiel von Fenstergröße und Höhenschwelle.

Zum Laufzeitverhalten ist generell zu sagen, dass bei der Anwendung der Grundinterpolation auf kleine Szenen die Bearbeitungszeit im Wesentlichen durch die manuelle Interaktion geprägt ist. Erst bei größeren Bildszenen spielt die Laufzeit des Minimumabgleichs eine Rolle. Bei der pyramidalen Verfeinerung der Grundhöhe hängt die Laufzeit bei großen Bildszenen stark von der Zahl der Pyramidenebenen ab. Maßgeblich sind die letzten Auflösungsebenen, also nahe der Originalauflösung. Messwerte zu einem konkreten Testbeispiel werden im folgenden Abschnitt diskutiert. In Tabelle 3.1 sind allgemein die Vor- und Nachteile der beiden Ansätze gegenübergestellt.

### 3.4.2 Qualitative und quantitative Bewertung

Zur Bewertung der Genauigkeit der vorgestellten Methoden sind Referenzmessungen nötig. In diesem Fall also Grundhöhen — auch in Regionen mit erhöhten Objekten. Prinzipiell kämen dafür topografische Karten oder Karten der Katasterämter in Betracht. Die Genauigkeit topografischer Karten kann jedoch nicht den Anforderungen an eine Referenzmessung genügen<sup>4</sup>. MÖLLER (2001) verwendet sowohl zur Bestimmung verlässlicher Grundpunkte für eine DOM-Interpolation wie auch zur abschließenden Qualitätsbeurteilung die Automatisierte Liegenschaftskarte (ALK). Dieses in Deutschland genaueste Kartenmaterial zur Erfassung von Liegenschaften enthält auch Angaben über Gebäudehöhen. Anhand dieser Höhenangaben erfolgt die Bewertung des abgeleiteten DTM.

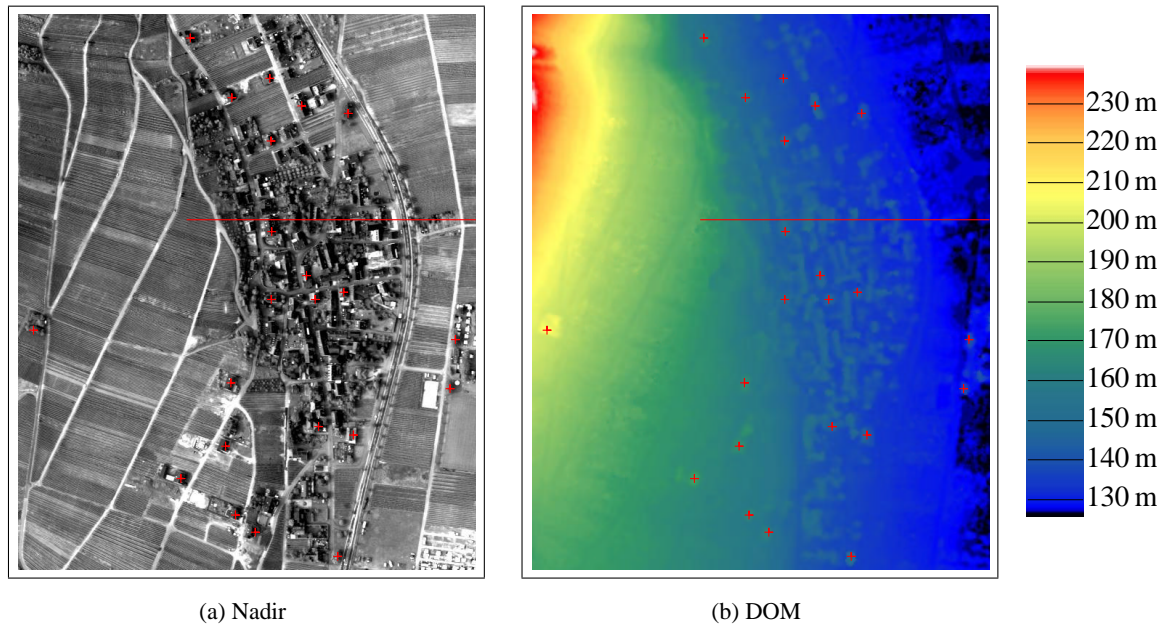
Weiterhin kämen die Daten eines Laserabtasters in Frage. Durch die Auswertung des jeweils zuletzt eintreffenden Echos („Last Pulse“) werden die Entfernung des am weitesten vorgedrungenen Laserstrahls erfasst. In bewaldeten Gebieten kann so auch die Grundhöhe ermittelt werden. Offensichtlich funktioniert dies nicht bei komplett für den Laserstrahl undurchdringlichen Objekten wie Häusern. Außerdem würden sich bei einer Gegenüberstellung mit Laser- oder Kartenreferenzdaten deren Ungenauigkeiten wie auch die Ungenauigkeit des HRSC-DOM in der Auswertung widerspiegeln.

Aus diesen Gründen werden in einem HRSC-DOM für mehrere erhöhte Punkte die Referenzgrundhöhen manuell bestimmt. Dies erfolgt durch Betrachtung und Mittelung der ein erhöhtes Objekt umgebenden Grundpunkte. Der gewählte Datensatz ist in Abbildung 3.7 dargestellt. Der  $632 \times 767$  Pixel messende Ausschnitt mit einem Meter geometrischer Auflösung zeigt ein Dorf in einer Weinbauregion. Während der rechte Bildbereich flach ist, erhöht sich das Gelände zur linken oberen Bildecke hin. An 21 in den Abbildungen durch ein rotes Kreuz markierten Stellen wird die Grundhöhe in der obengenannten Weise manuell bestimmt.

Mit dem Grundinterpolationsverfahren gelangt man mit folgender Parameterkonstellation schnell zu einem zufriedenstellenden Segmentierungsergebnis (erhöhte/nicht-erhöhte Punkte): Minimumabgleich durch ein Fenster mit 20 m Kantenlänge und 3 m Höhengschwelle sowie durch ein Fenster mit 5 m Kantenlänge und 1 m Höhengschwelle. Die Segmentierungsergebnisse werden anhand eines kreisförmigen strukturierenden Elements mit 5 Pixel Durchmesser dilatiert und anschließend logisch UND-verknüpft. Die Grundhöhe der erhöhten Regionen wird dann, wie in Abschnitt 3.2 beschrieben, interpoliert.

Gute Parameter zur pyramidalen Verfeinerung der Grundhöhe werden für diese Auswertung wie folgt ermittelt: Die aus der Grundinterpolationsmethode stammende Segmentierung in erhöhte Punkte und Grund wird als Referenz für eine automatisierte Parameterwahl herangezogen. Die Höhendifferenz von DOM und pyramidal verfeinertem DTM soll für die als Grund identifizierten Punkte möglichst gering ausfallen. In mehreren Durchläufen werden folgende Parameter in einem sinnvoll

<sup>4</sup>Die von den deutschen Landesvermessungsämtern angebotenen Geländemodelle weisen je nach Produkt Rasterweiten von 10–50 m mit Höhenggenauigkeiten von 0,5–5 m auf.

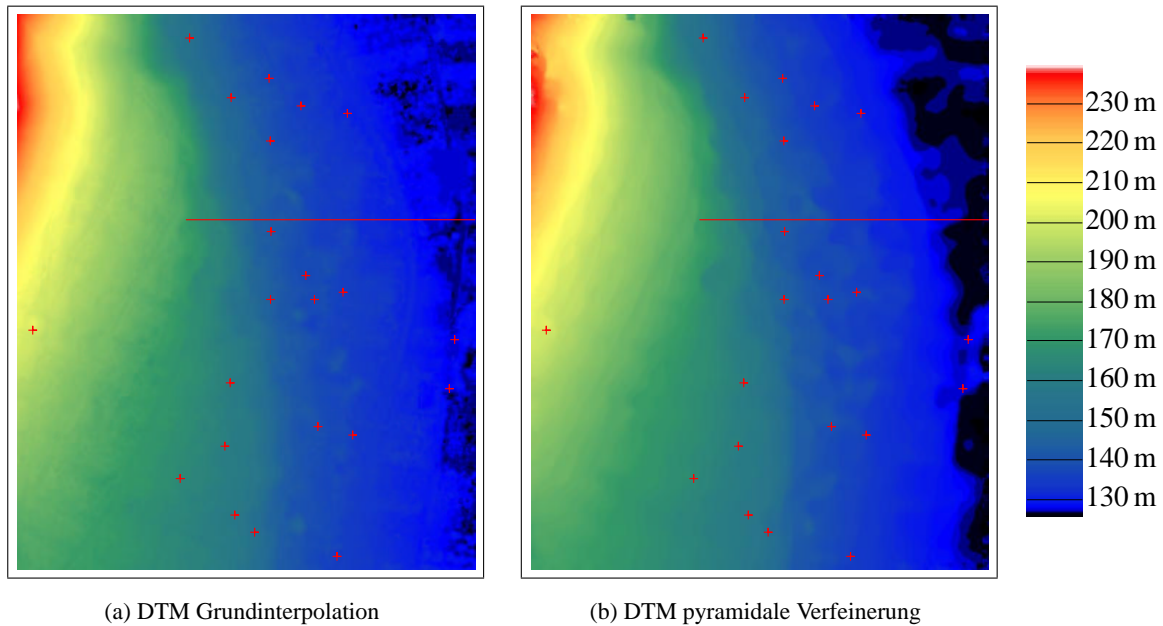


**Abbildung 3.7:** Testdatensatz zur Beurteilung der Methoden zur DOM-Normierung. In (a) ist zur Verdeutlichung der Nadirkanal dargestellt, in (b) das zugehörige farbkodierte DOM.

erscheinenden Wertebereich variiert: Die Zahl der Pyramidenebenen (5, 6 und 7), der Skalierungsfaktor zwischen den Pyramidenebenen (1,5, 1,7 und 2,0), der Fensterskalierungsfaktor relativ dazu (0,9, 1,0, 1,1 und 1,2) sowie die maximal zulässige Geländeneigung (0,7, 0,8, 0,9 und 1). Mit der Konstellation 7 - 1,7 - 1,1 - 0,9 kann die geringste mittlere und maximale Höhenabweichung zu den Grundpunkten im DOM erzielt werden. Das resultierende Terrainmodell wird anschließend mit einem  $21 \times 21$  Medianfilter geglättet.

#### Qualitative Beurteilung

Die so resultierenden Geländemodelle sind in Abbildung 3.8 einander gegenübergestellt. Wie man sieht, ist mit beiden Herangehensweisen die Geländemorphologie gut erfasst worden. Zur Veranschaulichung von Unterschieden wird entlang der rot eingezeichneten Linie in den beiden Terrainmodellen wie auch im DOM ein Höhenprofil abgegriffen (s. Abb. 3.9). Die Schnittlinie enthält mehrere erhöhte Objekte. Deutlich zu erkennen ist das lineare Interpolationsverhalten der Grundinterpolationsmethode (z. B. im Abszissenbereich 130–190 m). Demgegenüber ist die Grundlinie bei erhöhten Objekten im iterativ verfeinerten DTM deutlich nichtlinear. Die dort — innerhalb des Bereichs eines erhöhten Objekts — vorkommenden Sprünge im Terrain erscheinen unrealistisch. Aus Mangel an besserem Wissen scheint eine lineare Interpolation die bessere Wahl zu sein. Bei der Planung von Gebäuden bezieht sich die Gebäudehöhe jedoch üblicherweise auf die Höhe der Kelleroberkante. Würde dies in einem DTM berücksichtigt werden, müsste die Grundfläche dort auch



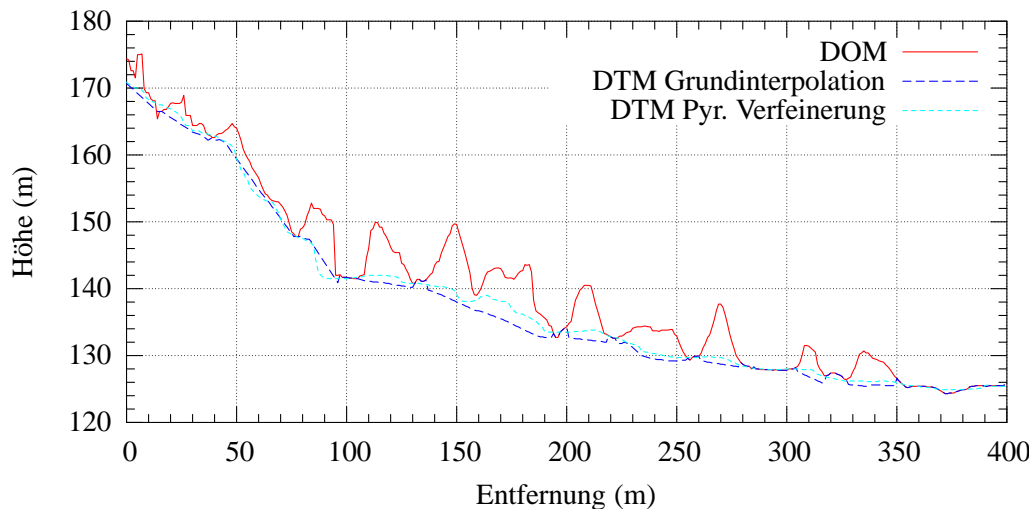
**Abbildung 3.8:** Farbgekodierte Darstellung der Terrainmodelle, die durch Grundinterpolation (a) und pyramidale Verfeinerung der Grundhöhe (b) gewonnen wurden. Die Parameter zur Erzeugung der Terrainmodelle werden im Text besprochen.

in Hanglage eben erscheinen.

Bei der Position 160 m im Höhenprofil ist zu erkennen, dass die Vertiefung durch das Grundinterpolationsverfahren nicht als Grund eingeordnet wird. Dies kann durch die Dilatation der Nicht-Grundsegmentierung bedingt sein. Resultat ist eine (ungenauere) Grundinterpolation von den nächstgelegenen Grundpunkten aus.

### Genauigkeit

Misst man die Höhenabweichung der generierten Terrainmodelle zu den Vorgaben an den manuell vermessenen erhöhten Punkten, erhält man das in Tabelle 3.2 aufgeführte Resultat. Die mittleren Abweichungen halten sich in einem für diese Arbeit gut vertretbaren Rahmen (0,43 m bzw. 1,09 m mit relativ geringen Standardabweichungen). Die Messung der Höhenabweichung gibt den in mehreren Experimenten gewonnenen Eindruck wieder, dass der Ansatz der pyramidalen Verfeinerung zu ungenaueren Ergebnissen führt als die Grundinterpolation. Zu beachten ist bei den Angaben zur Genauigkeit der beiden Methoden, dass die Messung nur erhöhte Punkte beinhaltet. Da die Höhenabweichung bei realen Grundpunkten für beide Ansätze sehr gering ist, ergeben sich über das gesamte Geländemodell gesehen offensichtlich wesentlich höhere Genauigkeiten.



**Abbildung 3.9:** Gegenüberstellung eines DOM-Höhenprofils mit Profilen von Terrainhöhen aus Grundinterpolation und pyramidaler Verfeinerung der Grundhöhe. Die Höhenwerte wurden entlang der in den Abb. 3.7 und 3.8 rot eingezeichneten Linie abgegriffen.

	Grundinterpolation	Pyramidale Verfeinerung
Mittlere Abweichung (m)	0,43	1,09
Standardabweichung (m)	0,37	0,61
Maximale Abweichung (m)	1,50	2,10

**Tabelle 3.2:** Genauigkeitsangaben für die beiden vorgestellten Methoden zur DOM-Normierung. Grundlage sind die in Abb. 3.7 durch rote Kreuze dargestellte Referenzpunkte, für die manuell die Grundhöhe ermittelt wurde.

#### Rechenzeit

Der eingangs erwähnte Testdatensatz wird auch herangezogen, um die Laufzeiten der beiden Verfahren einander gegenüber zu stellen. Wie bereits erwähnt, zeichnet sich der Ansatz zur Grundinterpolation durch einen höheren Anteil an manueller Interaktion aus. Dadurch wird die Verarbeitungszeit bei kleineren Szenen maßgeblich geprägt. In Tabelle 3.3 sind die reinen Rechenzeiten für die beiden Verfahren aufgeführt. Die Angaben beziehen sich auf das in Abschnitt 1.5.1 beschriebene Rechnersystem. Der erste Teil der Grundinterpolation, der Minimumabgleich mit anschließender Schwellwert- und morphologischer Öffnungsoperation, benötigt je nach Größe des Abgleichfensters im Mittel 21,9 bzw. 19,6 s (20 bzw. 5 m Kantenlänge des Fensters). Die Zeit zur Kombination der beiden Segmentierungsergebnisse ist zu vernachlässigen. Der Interpolationsvorgang inklusive Triangulation nimmt 5,7 s in Anspruch. Insgesamt ergibt sich somit eine Rechenzeit von etwa 47 s, was deutlich unter der manuellen Interaktionszeit von etwa 2–3 Minuten liegt. Eine zufriedenstellende Parameterkonstellation vorausgesetzt, ist bei der pyramidalen Verfeinerung der Geländehöhe die Pro-



grammlaufzeit gleich der Bearbeitungszeit für die Szene. Mit den angegebenen Parametern beträgt die Rechenzeit 163 s. Wird nur die reine Rechenzeit betrachtet, ist das Grundinterpolationsverfahren somit wesentlich schneller. Durch die bei einem realistischen Einsatz nicht zu vernachlässigende manuelle Interaktionszeit liegen die beiden Verfahren jedoch etwa gleich auf.

Grundinterpolation (s)	Pyramidale Verfeinerung (s)
Normierung, Schwellwert & Morph. Öffnen	
Fenster mit 20 m Kantenlänge	21,9
Fenster mit 5 m Kantenlänge	19,6
Interpolation	5,7
$\Sigma$	47,2
	163,0

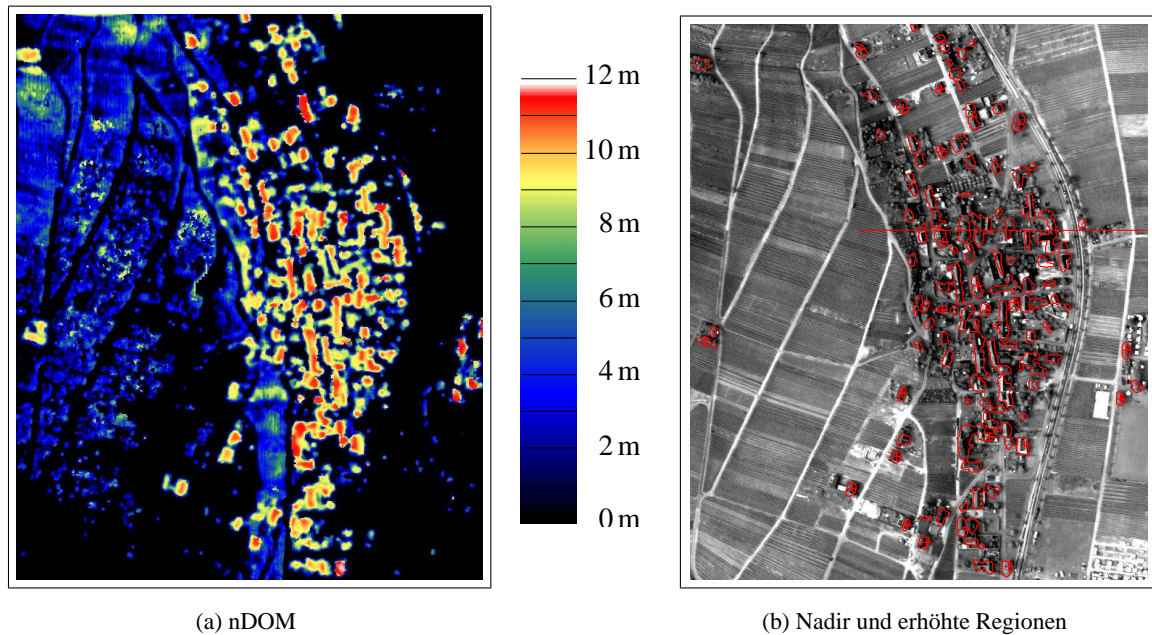
**Tabelle 3.3:** Laufzeitangaben für die beiden vorgestellten Methoden zur Ableitung eines DTM aus einem DOM. Grundlage ist das in Abb. 3.7(b) dargestellte DOM mit  $632 \times 767$  Pixeln.

### 3.5 Objektextraktion

Nachdem in den Abschnitten 3.2 und 3.3 beschrieben wurde, wie aus einem Oberflächenmodell ein Geländemodell abgeleitet werden kann, wird nun auf die verbleibenden Schritte zur Extraktion erhöhter Objekte eingegangen. Durch die Subtraktion von DOM und daraus abgeleitetem DTM ergibt sich ein normiertes Oberflächenmodell, nDOM. In Abbildung 3.10(a) ist das aus dem Testdatensatz des letzten Abschnitts und dem Grundinterpolationsverfahren hervorgehende nDOM dargestellt. Wie zu erwarten, treten die erhöhten Objekte nun deutlich hervor. Mittels einer Schwellwertoperation können erhöhte Segmente leicht identifiziert werden. Jedoch ist darauf zu achten, die Hörschwelle nicht zu tief anzusetzen. Da weder die Generierung des Ausgangs-DOM noch das Verfahren zur Terrainbestimmung fehlerfrei sind, wird ein geringfügig über der Nullhöhe liegender Schwellwert auch das Fehlerrauschen dieser beiden Prozesse erfassen. Eine zu hoch angesetzte Hörschwelle hingegen kann bei weniger hohen Objekten zu einer Konturlinie führen, die das eigentliche Objekt nicht mehr befriedigend repräsentiert.

Wie in Abschnitt 1.4.2 diskutiert, spielt neben den eigentlichen Pixelwerten, hier also Höhenangaben der jeweiligen Region, bei der Objektbeschreibung die Kontur des Objekts eine wesentliche Rolle. Die Konturlinie gibt dabei die Grenze vom Objekt zur Umgebung an. Bei den traditionellen Konturverfolgungsalgorithmen wird eben diese Grenze (genauer: Pixel des betrachteten Segments, die Nicht-Segmentpixel als Nachbarn haben), ermittelt (PAVLIDIS 1982, S. 143–146). Im Fall erhöhter Objekte und Vorhandensein eines nDOM wird die binäre Objekt/Nicht-Objekt Segmentierung durch den bereits genannten Hörschwelle erreicht.

Beginnend von einem Startpunkt der Konturlinie wird diese iterativ verfolgt durch Überprüfen



**Abbildung 3.10:** Farbodierte Darstellung des durch Grundinterpolation normierten Oberflächenmodells (a), sowie der zugehörige Bilddausschnitt überlagert mit den rot dargestellten Konturlinien der erhöhten Regionen (b).

aller vier bzw. acht Nachbarn des aktuellen Punkts, Fortschreiten zum nächsten Punkt und Einhalten einer festen Rangfolge für die Suchrichtungen. Eine Kodierung der Konturlinie in Form der Richtungen der Verfolgungsschritte („Kettencode“) bietet sich an.

Insbesondere zur Konturverfolgung in nicht-binären Segmentierungen haben sich neben der traditionellen Definition der Konturlinie eine erweiterte Definition (s. FENG UND PAVILIDS 1975, PAVLIDIS 1977, PAVLIDIS UND LIOW 1990) und darauf aufbauende Verfolgungsalgorithmen (s. z. B. LIOW 1991) durchgesetzt. Wesentlicher Vorteil dieser Konturliniendefinition ist, dass zwei direkt benachbarte Regionen nun eine gemeinsame Konturlinie besitzen. Die konsistente Darstellung der Segmentkontur erleichtert nachfolgende Auswerteprozesse.

In Einzelfällen auftretende Probleme der Konturverfolgungsverfahren sind der Verlust innerer Konturen, der Verlust einer durchgehenden Verbindung sowie das mehrmalige Verfolgen einer Konturlinie. Bei Verwendung der erweiterten Konturliniendefinition stimmt zudem die reproduzierte Konturlinie nicht mit der Originalkonturlinie überein. Zu einer effizienten Lösung dieser Probleme sei auf den Ansatz von REN ET AL. (2002) verwiesen.

Die an dieser Stelle thematisierte Konturverfolgung kommt immer dann zum Einsatz, wenn Objekte nicht mehr pixelbasiert, sondern vektoriell repräsentiert werden sollen (s. auch Abschnitt 1.4.2). Dies wäre hier der Fall, wenn eine Umrissbeschreibung der erhöhten Objekte gewünscht ist. Aber auch in allen nachfolgenden Verarbeitungsschritten ist der Übergang zu einer vektoriellen Darstel-

lung möglich. Für die in Kapitel 6 behandelte Verbesserung der Gebäudekonturen stellen Umrissspolygone sogar die Arbeitsgrundlage dar. Außerdem werden in jedem der nachfolgenden Verarbeitungskapitel die Ergebnisse immer auch in Form von überlagerten Konturpolygonen visualisiert.

Falls Größenattribute über die betrachteten Objekte vorliegen, können diese bereits vor der Konturverfolgung überprüft werden. Im Hinblick auf eine Gebäudeextraktion können Segmente, die kleinere Abmessungen als für ein Haus zulässig aufweisen, ausgeschlossen werden. Abbildung 3.10(b) zeigt die Konturen von Regionen, die mindestens 50 Pixel (entspricht  $50\text{ m}^2$ ) beinhalten und höher als 5 m sind.



## 4 Segmentierung erhöhter Objektregionen

Das Ziel dieses Kapitels ist eine Differenzierung der durch die Methoden des vorhergehenden Kapitels gewonnenen erhöhten Regionen. Die zwei maßgeblichen Kategorien erhöhter Objekte im städtischen wie auch im ländlichen Gebiet sind Gebäude und Bäume. Aus diesem Grund wird in diesem Kapitel vereinfachend davon ausgegangen, dass erhöhte Objekte einer dieser beiden Kategorien angehören. Die Detektion von Gebäuden und Bäumen reduziert sich somit auf die Unterteilung der erhöhten Regionen in diese beiden Klassen.

Gebäudematerialien und Vegetation lassen sich anhand ihrer spektralen Charakteristika gut unterscheiden. Daher kommen die Multispektraleigenschaften der HRSC für die in diesem Kapitel vorgestellten Methoden voll zum Tragen. Der Anstieg der Reflektanz beim Übergang vom roten in den infraroten Spektralbereich ist bei Vegetation deutlich stärker als bei versiegelten Flächen. Deshalb werden bei dem in diesem Kapitel vorgeschlagenen Vorgehen Vegetationsbereiche aufgrund dieser charakteristischen Eigenschaft detektiert. Baumregionen resultieren dann aus der Verschneidung der Vegetationssegmente mit den erhöhten Regionen. Die Gebäudesegmente wiederum ergeben sich initial aus den verbleibenden erhöhten Regionen. Dieser Thematik widmet sich Abschnitt 4.2, der zentrale Teil dieses Kapitels.

Die initialen Gebäudesegmente werden, wie in Abschnitt 4.3 erörtert, in weiteren Verarbeitungsschritten überarbeitet. Einerseits werden Schattenbereiche ausgeschlossen, andererseits wird die Segmentform nachbearbeitet.

Die vorgestellten Algorithmen werden in Abschnitt 4.4 im Hinblick auf ihre Genauigkeit an einem Testdatensatz ausgewertet.

### 4.1 Einführung und Literaturüberblick

Die Unterscheidung von Gebäuden und Bäumen kann anhand unterschiedlicher Ausgangsdaten erfolgen. Die Ansätze von DISSARD ET AL. (1997) und ECKSTEIN UND MUNKELT (1995) bauen ausschließlich auf Grauwert- bzw. multispektralen Merkmalen auf.

Die Grauwerttextur (Entropie) bereits vorliegender Segmente ist bei DISSARD ET AL. (1997) Grundlage für einen Regionenwachstumsprozess. Resultat ist eine neue Unterteilung der Segmente. Diese neuen Segmente werden abschließend in Gebäude und Vegetation klassifiziert.

Einfache Grauwert- oder multispektrale *Texturmerkmale* verwenden auch ECKSTEIN UND MUNKELT (1995), um erhöhte Regionen in Bäume und Häuser zu klassifizieren.

Bei der Bestimmung textueller Merkmale ist das einleitend (Abschnitt 1.1, S. 5) im Zusammenhang mit einer objektbasierten Interpretation diskutierte Problem der *Objektgrenzen* zu beachten. In den zwei gerade erwähnten Ansätzen werden die Texturmaße fensterbasiert ermittelt. Eine genaue Ermittlung der Objektgrenzen wird so sehr schwierig.

Eine sichere, nur auf spektralen Charakteristika aufbauende Klassifikation in verschiedene Oberflächenmaterialien ist nur bei einer hinreichend hohen spektralen Auflösung möglich. Deswegen nutzen beispielsweise FORD ET AL. (1997) hyperspektrale Bilddaten für eine pixelweise Klassifikation in unterschiedliche Materialtypen. Bei Fernerkundungssensoren mit hoher spektraler Auflösung ist die räumliche Auflösung üblicherweise geringer. Die entstehenden „Mischpixel“ können nur durch aufwändige Verfahren näherungsweise in die einzelnen Objektkategorien separiert werden. Detaillierte Objektstrukturen lassen sich so nur schwer extrahieren.

Ist die spektrale Auflösung geringer, also im monochromatischen oder multispektralen Fall, ist die Einbeziehung von Höheninformation sinnvoll.

Bei HAALA UND BRENNER (1999) fließt die Höheninformation eines normierten Höhenmodells zusammen mit den drei Farbkanälen eines Colorinfrarotbildes (CIR) direkt in eine unüberwachte Klassifikation ein. Die dadurch gewonnene Vorsegmentierung wird *manuell interpretiert und kombiniert*, um zu einer Klassifikation in Straßen, Gras, Bäume, Gebäude und Schatten zu gelangen. Neben der manuellen Überarbeitung der Klassifikationsergebnisse benötigt der unüberwachte Klassifikationsschritt *abstrakte Vorgaben*. So müssen die Anzahl der Klassen für die Vorsegmentierung oder die maximal zulässige Standardabweichung der Merkmale innerhalb eines Segments definiert werden.

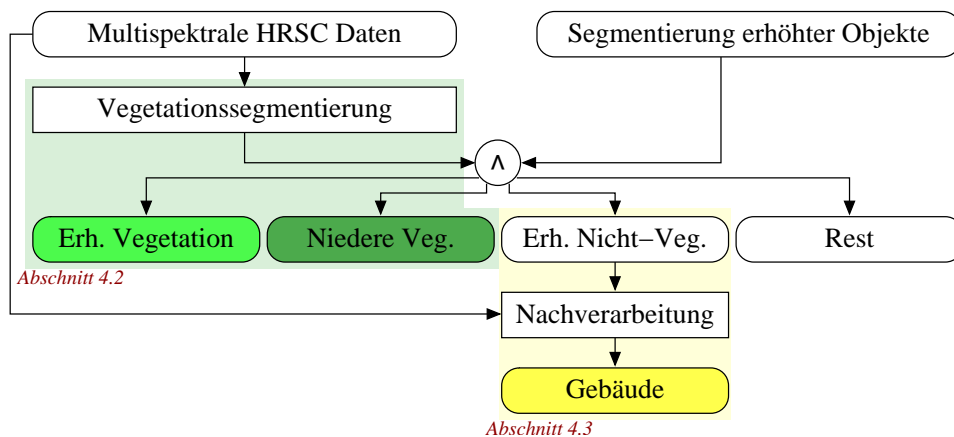
BRUNN (2001) nutzt ein Bayes'sches Netz, um mittels Höhen- und Farbinformation Gebäude zu detektieren<sup>1</sup>. Bäume sind in diesem auf die Rekonstruktion von Gebäuden ausgerichteten Ansatz kein Untersuchungsgegenstand, sie werden implizit durch die Gebäudedetektion ausgeschlossen.

GIRARD ET AL. (1998) verwenden zur Gebäudedetektion neben einem Oberflächenmodell die Farbinformation aus einem RGB-Bild. Dabei werden *Gebäude durch entsprechende Merkmale direkt charakterisiert*. Sie müssen einen roten Farbton, eine hohe Farbsättigung und eine entsprechende Höhe aufweisen. Aufgrund unterschiedlicher Materialien sind Hausdächer jedoch spektral sehr heterogen. Eine umfassende Modellierung ihrer spektralen Charakteristika ist nur anhand der bereits erwähnten spektral höher auflösenden Bildprodukte möglich.

Ausschließlich auf Höheninformation baut der Ansatz von BRUNN UND WEIDNER (1997) auf. Sie schlagen vor, die Varianz der Oberflächennormalen eines DOM zur Vegetationsdetektion zu nutzen. Bei Bäumen ist die Oberfläche im Höhenmodell rau, somit ergibt sich eine hohe Varianz; bei künstlich errichteten Objekten ist die Oberfläche im Allgemeinen glatt, es resultieren niedrige Varianzwerte. Auch CORD ET AL. (2001) nutzen die Richtung lokaler Oberflächennormalen eines DOM. Die Richtungen werden für jedes Segment in einem Histogramm gesammelt. Über einen Schwell-

---

<sup>1</sup>Höhe über Grund aus einem nDOM und Varianz der Oberflächennormalen. Erläuterungen zu Letzterem folgen in den Angaben zu BRUNN UND WEIDNER (1997).



**Abbildung 4.1:** Schematische Darstellung zur Segmentierung in Vegetation und Gebäude.

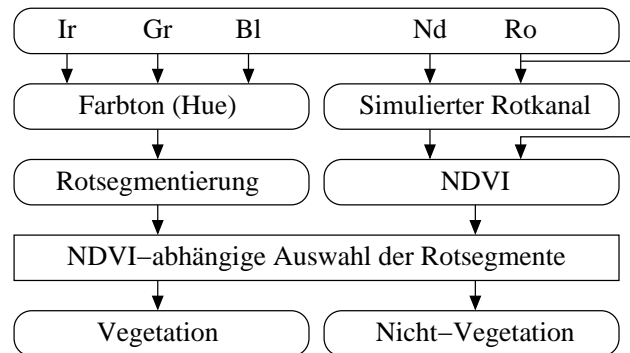
wert werden Gebäude- und Vegetationssegmente unterschieden.

Wird nur die Höheninformation zur Unterscheidung von Bäumen und Gebäuden benutzt, muss ein *sehr präzises Höhenmodell* vorliegen. Außerdem sollten nur die tatsächlich gemessenen Höhenpunkte verwendet werden. Ein zu einem vollständigen Raster interpoliertes Oberflächenmodell kann zu einer vermeintlichen Glätte führen (s. auch Abschnitt 2.3.5).

Wie der Großteil der vorgestellten Ansätze nutzt auch die in dieser Arbeit gewählte Strategie zur Unterscheidung von Bäumen und Gebäuden das Potenzial der *Kombination von Höheninformation und multispektraler Information*. Abbildung 4.1 gibt einen schematischen Überblick über das *zweistufige Vorgehen*. In einem ersten Schritt werden Vegetationsregionen durch einen auf den Multispektraldaten aufbauenden Segmentierungsprozess identifiziert. Da die *Vegetationssegmentierung* schon seit mehreren Jahrzehnten Gegenstand der Erdfernerkundung ist, kann diesbezüglich auf ein breites Fundament an Ansätzen und Untersuchungen zurückgegriffen werden. Außerdem lässt sich grüne, vitale Vegetation trotz ihrer mannigfaltigen Erscheinungsformen gut anhand der spektralen Zusammensetzung des reflektierten Lichts charakterisieren. Dazu genügt eine „multispektrale Auflösung“. Gebäudedächer sind aufgrund der damit verbundenen Materialvielfalt auch im spektralen Merkmalsraum weniger konzentriert verteilt und somit schwieriger zu identifizieren. Eine auf einer vorgeschalteten Vegetationssegmentierung beruhende Gebäudedetektion umgeht dieses Problem. Es werden keine explizit formulierten Anforderungen an die spektrale Zusammensetzung des reflektierten Signals der Gebäudesegmente gestellt.

Aus der Kombination mit einer aus einem nDOM abgeleiteten Segmentierung erhöhter Objekte ergeben sich Segmente erhöhter und niederer Vegetation (Baumgruppen / Wiese, Büsche, Rasen). Die erhöhten Nicht-Vegetationssegmente werden in einem zweiten Schritt zu Gebäudesegmenten nachverarbeitet.

Die Vegetationssegmentierung erfolgt durch *pixelweises* Betrachten von Segmentierungskriterien. Anders als bei einem fensterbasierten Vorgehen, beispielsweise der Bestimmung textueller Merk-



**Abbildung 4.2:** Schematische Darstellung zur Vegetationssegmentierung.

male, können die Grenzen der Vegetationssegmente so genau bestimmt werden. Die Kombination mit der Segmentierung erhöhter Objekte aus einem fensterbasiert bestimmten nDOM (bzw. aus einem mit Ungenauigkeiten behafteten DOM) führt freilich ebenfalls zu ungenaueren Objektgrenzen. Für die Gebäudesegmente folgt deshalb in Kapitel 6 eine Überarbeitung der Konturlinien.

Die Segmentierungskriterien sind *konkrete spektrale Merkmale*, die dem Nutzer auch als Zwischenprodukte zur Verfügung stehen. Das Resultat der durch Schwellwerte bestimmten Segmentierung ist so auch für Bildverarbeitungslaien nachvollziehbar. Die Segmentierungsergebnisse sind *bedeutungstragend* (z. B. „Gebäude“) und bedürfen keiner weiteren Interpretation. Die manuelle Interaktion ist auf die Feinabstimmung der Schwellwerte beschränkt.

## 4.2 Vegetationssegmentierung

Wie eingangs erwähnt, erfolgt die Vegetationssegmentierung ausschließlich aufgrund multispektraler Bildinformation. Die Multispektralkanäle der verschiedenen HRSC-Sensoren unterscheiden sich jedoch in Bezug auf deren spektralen Empfindlichkeitsbereich (s. Kapitel 2). Dies hat für das vorgeschlagene Vorgehen zur Vegetationssegmentierung zusätzliche Arbeitsschritte bei Nutzung von HRSC-A Daten zur Folge (s. Abbildung 4.2), auf die nachfolgend eingegangen wird. Die Unterschiede bei einem Einsatz von HRSC-AX Daten werden in Abschnitt 4.2.4 diskutiert.

Aus der multispektralen Information wird ein in der Erdfernerkundung üblicher Vegetationsindex berechnet (NDVI: „Normalized Difference Vegetation Index“). Durch diesen Index werden reine Vegetationspixel durch entsprechend hohe Werte charakterisiert.

Um eine Segmentform, die Vegetationsbereiche möglichst zusammenhängend erfasst, zu erzielen, wird ein weiteres Kriterium hinzugezogen. Anhand des in einer bestimmten Kanalkombination für Vegetation typischen roten Farbtons erfolgt eine Bildsegmentierung. Die resultierenden Regionen werden durch den bereits erwähnten Vegetationsindex in einem zweiten Schritt verifiziert. Kombiniert man die so entstandenen Vegetationssegmente mit der im vorhergehenden Kapitel eingeführten



Segmentierung erhöhter Regionen, ergeben sich erhöhte Vegetationssegmente. Damit sind die Baumregionen identifiziert.

#### 4.2.1 Der Vegetationsindex NDVI

In der Erdfernerkundung dient die spektrale Signatur als wesentliches Unterscheidungsmerkmal von Vegetation und sonstigen Gegenständen. Die Blätter aktiver, grüner Vegetation weisen einerseits wasserreiche Zellinhalte und andererseits interzelluläre Lufträume auf. Durch die damit verbundenen großen Unterschiede in den Brechungsindizes sind Blätter gute Lichtstreuer (s. RAY 1994). Durch die Blattpigmente (wie Chlorophyll, Protochlorophyll, Xanthophyll) wird das Licht im sichtbaren Bereich des Spektrums (400–700 nm) stark absorbiert; die Vegetation erscheint also dunkel. Die geringste Absorption in diesem Wellenlängenbereich ist um 550 nm; deshalb erscheinen uns die Blätter grün. Die Elektronenübergänge, die die Absorption im sichtbaren Bereich verursachen, können von längerwelliger Strahlung nicht mehr hervorgerufen werden. Auch die Molekularbewegung, die die Absorption langwelliger Strahlung (1 300–2 500 nm) bedingt, ist für Strahlung im nahen Infrarot (700–1 300 nm) unwirksam. Daher gibt es im nahen Infrarot keine starke Absorption; Vegetation streut das Licht, wie eingangs geschildert, hier sehr gut.

Der sprunghafte Anstieg des Reflexionsgrades um 700 nm ist charakteristisch für gesunde Vegetation. Unbewachsener Boden hingegen weist, je nach Art, einen stetigen, geradlinigen Verlauf der Reflektanz auf. Je aktiver die Blattpigmente der Pflanzen sind, desto größer ist der Anstieg des Reflexionsgrades im nahen Infrarot. Zur Veranschaulichung dieses Zusammenhangs ist in Abbildung 4.3 exemplarisch das Reflektanzverhalten von Gras, trockenem Gras und Lehmboden einander gegenüber gestellt.

Aufgrund der geringen spektralen Auflösung von Multispektralscannern ist es nicht möglich, die spektrale Signatur z. B. von Vegetation direkt zu erfassen. Dafür kommen in der Erdfernerkundung spektral höher auflösende Systeme, die abbildenden Spektrometer (auch „Hyperspektralsensoren“), zum Einsatz. Bei der HRSC käme erschwerend hinzu, dass die Kanäle nur relativ zueinander, aber nicht absolut radiometrisch kalibriert sind (s. Abschnitt 2.3.2).

In Form von Vegetationsindizes hingegen wird versucht, einzelne spektrale Charakteristika der Vegetation zu erfassen. Insbesondere der für Vegetation typische Anstieg der Reflektanz um 700 nm dient als verbreitetes Merkmal. Der Anstieg wird üblicherweise anhand zweier die Schwelle von 700 nm umgebenden Spektralkanäle eines Erdfernerkundungssensors bestimmt. Ein Band erfasst die hohe Reflektanz im Nahen Infrarot, ein anderes die geringe Reflektanz im sichtbaren Bereich<sup>2</sup>. Eine starke Änderung der Reflektanz wird dann durch Differenzbildung ermittelt. Das erwähnte Reflexionsmaximum um 550 nm kann ebenfalls als Vegetationsmerkmal herangezogen werden. Jedoch wird von TUCKER (1979) gezeigt, dass die Kombination von Nah-Infrarot- und Rotkanal einer Kombina-

<sup>2</sup>Meist wird der Bereich um 650 nm gewählt, da dieser das Absorptionsmaximum innerhalb des üblicherweise in der Erdfernerkundung gemessenen Spektrums darstellt.

tion von Grün- und Rotkanal vorzuziehen ist.

Um unabhängiger von Albedoeffekten und Schatten zu werden, wird der Vegetationsindex NDVI („Normalized Difference Vegetation Index“) durch Ratiobildung normiert (ROUSE ET AL. 1973)<sup>3</sup>. Der Index ergibt sich als

$$\text{NDVI} = \frac{\text{NIR} - \text{ROT}}{\text{NIR} + \text{ROT}}. \quad (4.1)$$

Dabei seien mit ROT das den sichtbaren roten und mit NIR das den infraroten erfassende Band bezeichnet. Der Index bewegt sich in einem Wertebereich von  $-1, \dots, 1$ . Die Reflektanzkurve beispielsweise von Lehm Boden hat in dem betrachteten Wellenlängenbereich in etwa die Steigung 1. Im NDVI wird diese Steigung in etwa auf das Indexmittel 0 abgebildet. Bei aktiver, grüner Biomasse hingegen ist der Unterschied von NIR zu ROT höher, was in Werten größer als 0 resultiert. Als Richtwert<sup>4</sup> für die Untergrenze bei Vegetation gilt 0,3.

Erfolgreich angewendet werden kann der NDVI allerdings nur bei dichter Vegetationsbedeckung. Wird die Vegetation lichter, so mischen sich die spektrale Signatur von Hintergrund und Vegetation (ELVIDGE UND LYON 1985, HUETE ET AL. 1985). Bei der Nutzung geometrisch hochauflösender Sensoren verschiebt sich diese Mischpixelproblematik auf entsprechend kleinere Objekte und sollte sich folglich weniger stark auf die Detektion von Vegetationspixeln auswirken. Eine zeitversetzte Erfassung der Multispektralkanäle — wie es durch die unterschiedlichen Blickwinkel bei der HRSC der Fall ist — kann hingegen zu Überlagerungen verschiedener Objektteile führen. Ursache hierfür sind Bewegungen während der Aufnahme beispielsweise durch Wind (s. Abschnitt 2.5.1). Ein so zustande gekommener NDVI-Wert ist offensichtlich irreführend.

#### Die HRSC-A Rotkanäle

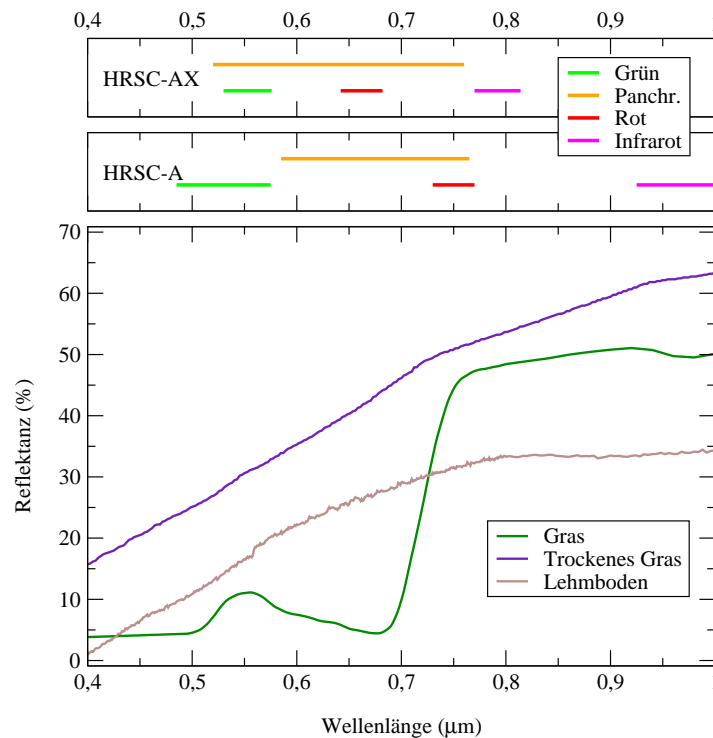
Wie in Kapitel 2 erwähnt wurde, ist die HRSC-A für den Einsatz im Mars-Orbit entworfen worden. Die Filter der Multispektralkanäle sind auf die Marsgeologie hin ausgelegt. Aus diesem Grund ist eine direkte Umsetzung der NDVI-Theorie für Daten der HRSC-A nicht möglich. Abbildung 4.3 verdeutlicht den Zusammenhang von Transmissionsbereich der HRSC-Filter und der Reflektanz gesunder Vegetation. Es ist gut zu erkennen, dass die HRSC-AX Generation auf Erdfernerkundungsanwendungen ausgerichtet ist. Der für grüne Vegetation charakteristische Wellenlängenbereich um 700 nm wird vom roten und infraroten Kanal eng flankiert.

Um die Kanäle der HRSC-A für NDVI-Ratios zu verwenden, müssen erst geeignete Rotkanäle abgeleitet werden. Der HRSC-A Rotkanal deckt den infrarotseitigen Bereich jenseits der 700 nm bereits relativ gut ab ( $750 \pm 20$  nm). Für obengenannten NDVI-Ratio wäre der HRSC-A Rotkanal also der Infrarotkanal. Der panchromatische Nadir liegt mit 675 nm zentraler Wellenlänge zwar unterhalb der 700 nm Schwelle, doch überdeckt er durch seine Breite von 180 nm den gesamten für Vegetati-

---

<sup>3</sup>Der unterschiedliche Einfluss der Atmosphäre auf NIR und ROT bleibt hierbei unbeachtet.

<sup>4</sup>Konkrete Werte hängen von vielen Faktoren ab, insbesondere von dem durch die gewählten Kanäle abgedeckten Spektralbereich.



**Abbildung 4.3:** Eine Gegenüberstellung des Transmissionsbereichs von HRSC-Filtern und der Reflektanz ausgewählter Materialien (die Spektren stammen aus der Aster Spektrenbibliothek (ASL)) im Bereich von 400–1 000 nm.

on charakteristischen Bereich. Aus diesem Grund muss der für NDVI-Ratios notwendige Rotkanal durch eine geeignete HRSC-Kanalkombination simuliert werden. Als Basiskanäle kommen dafür spektral nahe HRSC-A-Kanäle, hier also der panchromatische Nadir sowie der Rotkanal, in Frage.

Für eine vollständige Simulation eines künstlichen Rotkanals müssen neben dem Empfindlichkeitsbereich die Quanteneffektivität der CCD und die Transmission des optischen Systems berücksichtigt werden<sup>5</sup>. Eine grobe Abschätzung eines künstlichen HRSC-A Rotkanals wird meist anhand der Bandbreiten des roten und panchromatischen Kanals gewonnen. Auf dieser Grundlage leitet auch MÖLLER (2001) einen künstlichen Rotkanal ab. Wahrscheinlich<sup>6</sup> soll dort der Rotkanal entsprechend der Bandbreiten gewichtet werden, bevor er vom panchromatischen Kanal subtrahiert wird. GÄHLER (2000) zieht ebenfalls, allerdings ohne Gewichtung, den roten HRSC-Kanal vom panchromatischen ab. Jedoch wird festgestellt, dass der so abgeleitete künstliche Rotkanal nicht dem Spektralbereich des sichtbaren Rots entspricht. Ferner weist er kaum Informationsgehalt über die Vegetationsbedeckung auf und könne somit nur einen geringen Beitrag zur Vegetationsklassifikation leisten. Der von EHLERS ET AL. (2002) vorgestellte HRSC-Vegetationsindex baut ebenfalls auf dem so berechneten

<sup>5</sup> Alle genannten Parameter sind wellenlängenabhängig. Die jeweiligen Antwortfunktionen werden miteinander gefaltet, um die tatsächliche Empfindlichkeit der Optik-Filter-CCD-Kombination zu erhalten.

<sup>6</sup> Die angegebene Gleichung scheint fehlerhaft zu sein.

künstlichen Rotkanal auf. Mit einem Vegetationsindex aus der Differenz von künstlichem Rotkanal und Blaukanal oder aber von originalem Rotkanal und Blaukanal (jeweils HRSC-A Kanäle) werden dort die besten Ergebnisse erzielt.

Auch in dieser Arbeit erfolgt die Abschätzung eines künstlichen Rotkanals bandbreitenbasiert. Die so ermittelte Kanalgewichtung stellt einen groben Richtwert dar. Aufgrund der fehlenden absoluten radiometrischen Kalibration der HRSC-Kanäle ist eine Feinabstimmung der Gewichte bei unterschiedlichen Datensätzen ratsam. Zur NDVI-Ratiobildung wird infrarotseitig der HRSC-A Rotkanal (im Folgenden mit  $ROT_{org}$  bezeichnet) herangezogen. Um eine Gleichgewichtung der am NDVI beteiligten Kanäle zu erhalten wird der künstliche Rotkanal („ $ROT_{sim}$ “) mit gleicher Bandbreite und gleichem Abstand von der 700 nm-Marke wie der HRSC-A Rotkanal simuliert. Wie bereits erwähnt, deckt der panchromatische Kanal („PAN“) als einziger diesen Bereich des Spektrums ab, jedoch ist er zu breit. Die passende Linearkombination mit dem HRSC-A Rotkanal zur Subtraktion des Infrarotanteils kann durch eine äquivalente Kombination der Rotkanäle  $ROT_{org}$  und  $ROT_{sim}$  zur Simulation des panchromatischen Kanals berechnet werden. Da  $ROT_{sim}$  auf die gleiche Bandbreite wie  $ROT_{org}$  festgelegt ist, haben beide auch fast den gleichen Überlappungsanteil mit dem panchromatischen Kanal (s. Abbildung 4.4). Normiert auf die Summe eins sind die Anteile somit jeweils 0,5. Umgerechnet auf den künstlichen Rotkanal ergibt sich somit die Linearkombination  $2 \cdot PAN - ROT_{org}$ . Da der HRSC-A Rotkanal jedoch etwas weiter ins Infrarote reicht als der panchromatische Kanal ändert sich die Gewichtung geringfügig zugunsten des panchromatischen Kanals. Die Linearkombination zur Berechnung des panchromatischen Kanals ergibt sich wie folgt:

$$PAN = \frac{670 - 630}{765 - 585} ROT_{sim} + \frac{765 - 730}{765 - 585} ROT_{org} + \epsilon = \frac{40}{180} ROT_{sim} + \frac{35}{180} ROT_{org} + \epsilon \quad (4.2)$$

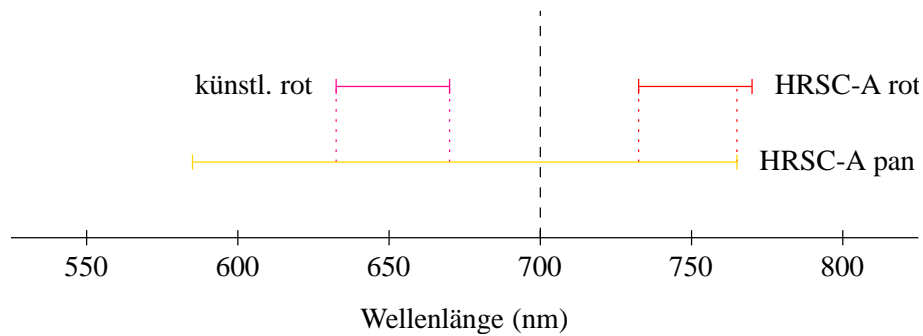
Wird der fehlende Anteil  $\epsilon$  außer Acht gelassen und die Summe der Gewichte auf eins normiert, so erhält man:

$$PAN = \frac{8}{15} ROT_{sim} + \frac{7}{15} ROT_{org} \quad (4.3)$$

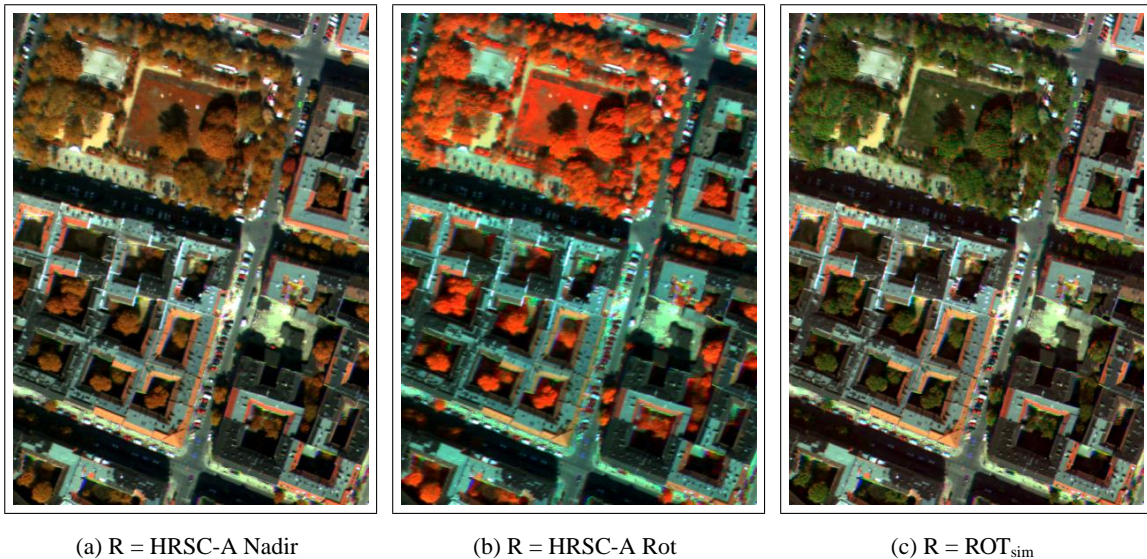
Daraus ergibt sich für den künstlichen Rotkanal folgende Linearkombination:

$$ROT_{sim} = \frac{15}{8} PAN - \frac{7}{8} ROT_{org} \quad (4.4)$$

Durch die Subtraktion vergrößert sich der Wertebereich im Vergleich zu den Ausgangskanälen und muss entsprechend angepasst werden. Zur Veranschaulichung des so gewonnenen künstlichen Rotkanals sind in Abbildung 4.5 drei RGB-Farbkomposite aus HRSC-A Daten dargestellt. Grün- und Blaukanal sind jeweils gleich und entsprechen dem grünen und blauen HRSC-Kanal. Nur der Rotkanal ist jeweils anders belegt. Einmal wird der panchromatische Nadirkanal (Abb. 4.5(a)), dann der im Infraroten liegende Rotkanal (Abb. 4.5(b)) sowie der aus den beiden Kanälen abgeleitete künstliche Rotkanal (Abb. 4.5(c)) verwendet. Insbesondere im Vergleich zur Verwendung des HRSC-A Rotka-



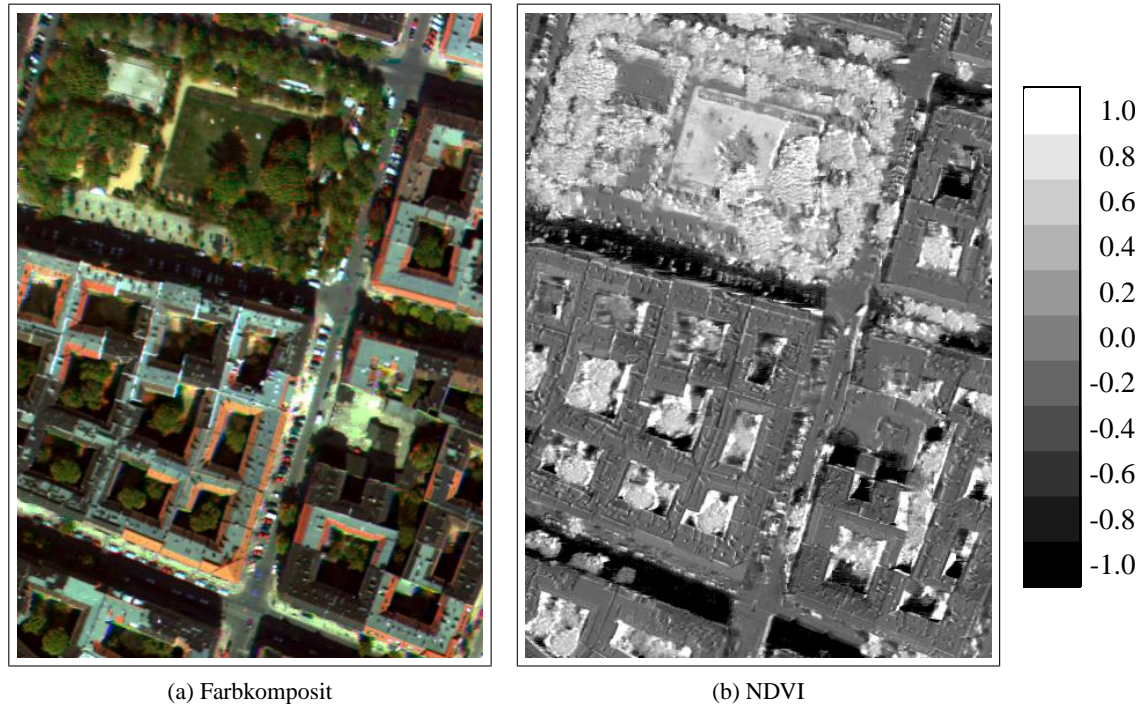
**Abbildung 4.4:** Schematische Darstellung der Bandbreiten des roten und panchromatischen HRSC-A Kanals sowie eines daraus abgeleiteten künstlichen Rotkanals.



**Abbildung 4.5:** Farbkomposite aus HRSC-A Daten und daraus abgeleitetem künstlichen Rotkanal. In der RGB-Darstellung sind der grüne und blaue Kanal jeweils mit dem entsprechenden HRSC-A Kanal belegt. Der Rotkanal ist in (a) der panchromatische HRSC-A Nadirkanal, in (b) der HRSC-A Rotkanal und in (c) der künstliche Rotkanal. Für diese Darstellung sind die Farbkanäle jeweils innerhalb des Grauwertbereichs mit einem Histogrammanteil von 2–98% linear gestreckt.

nals<sup>7</sup> führt der künstliche Rotkanal zu einem besseren Echtfarbeneindruck. Der künstliche Rotkanal entspricht also eher der menschlichen Wahrnehmung im roten Wellenlängenbereich.

<sup>7</sup>Durch den hohen Infrarotanteil des HRSC-A Rotkanals ergibt sich eine starke Reflektanz grüner Vegetation. Vegetation erscheint im Farbkomposit also rötlich.



**Abbildung 4.6:** HRSC-A Farbkomposit (a) (äquivalent zu Abb. 4.5(c)) und zugehörige Grauwertdarstellung des NDVI (b). Je höher der Vegetationsindex, desto heller ist hier der entsprechende Grauwert.

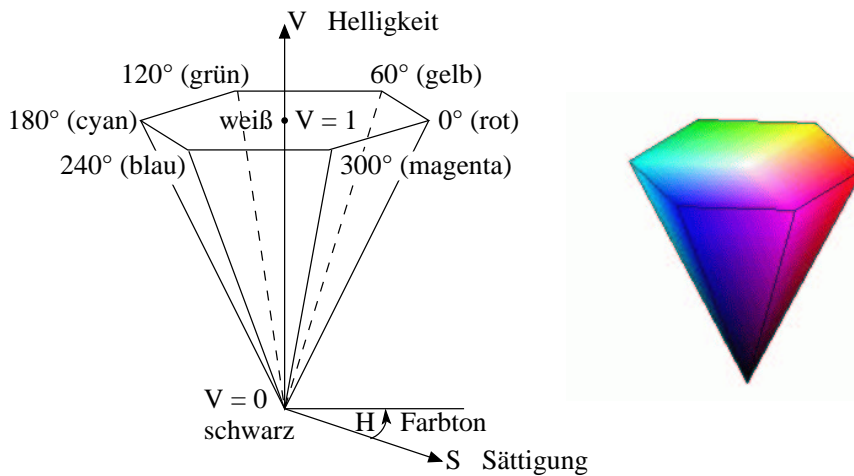
#### HRSC-NDVI

Nach der oben beschriebenen Methode wird nun aus den HRSC-Kanälen im roten Wellenlängenbereich — bei HRSC-A Daten nach Ableitung eines künstlichen Kanals — der Vegetationsindex NDVI berechnet. In Abbildung 4.6 ist der resultierende NDVI bildlich dargestellt. Grundlage dafür sind HRSC-A Daten einer Stadtscene mit Grünbereich. Es ist gut zu erkennen, dass Vegetation auch in Schattenbereichen auf hohe Indexwerte abgebildet wird. Gebäude, Straßen oder Sandflächen hingegen weisen meist negative Werte auf. Jedoch ist auch ersichtlich, dass Vegetation nicht auf einheitlich hohe Indexwerte abgebildet wird. Stellenweise sinken die NDVI-Werte auch z. B. innerhalb einer Baumgruppe deutlich ab. Dieser Effekt ist typisch für geometrisch hochauflösende Daten: Weniger dichte Baumkronen lassen immer wieder den Boden oder Äste erkennen. An diesen Stellen ist dann ein deutlich niedrigerer NDVI-Wert zu erwarten.

Im Auswertungsteil dieses Kapitels (Abschnitt 4.4.1) werden die theoretisch erreichbaren Fehleraten bei der Klassifikation in Gebäude und Bäume für Daten der HRSC-A und HRSC-AX einander gegenübergestellt. Dabei wird der Vorteil der HRSC-AX gegenüber der HRSC-A zur Vegetationssegmentierung quantitativ untermauert.

Mit Daten der HRSC-AX können alleine unter Nutzung des NDVI Gebäude und Bäume erfolg-





**Abbildung 4.7:** Schematische Darstellung des HSV-Farbraums.

reich getrennt werden. Wegen der für eine Vegetationssegmentierung eher ungeeigneten ausgerichteten Spektralkanäle der HRSC-A sollten für Daten dieses Sensors weitere Merkmale hinzugezogen werden. Im folgenden Abschnitt wird der Farbton als weiteres Merkmal für HRSC-A Daten eingeführt.

#### 4.2.2 Farbtonsegmentierung der Vegetation

Betrachtet man ein HRSC-Farbkomposit aus Infrarot-, Grün- und Blaukanal (RGB), so erscheint Vegetation in einem roten Farbton (s. z. B. Abbildung 4.5(b), dort allerdings mit dem im infraroten Bereich liegenden HRSC-A Rotkanal). In der Erdfernerkundung mit Colorinfrarotbildern (CIR) ist dieser rote Farbton ein typisches Vegetationsmerkmal. Im Wesentlichen basiert dieses — wie auch der zuvor beschriebene NDVI — auf einer für Vegetation typischen starken Reflektanz im nahen Infrarot und schwachen Reflektanz unterhalb 700 nm. In einer RGB-Darstellung, bei der das nahe Infrarot auf die rote Farbkomponente abgebildet wird, ist für Vegetation mit entsprechend roten Farbtönen zu rechnen. Sowohl bei der HRSC-AX wie auch bei der HRSC-A, deren Infrarotkanal wesentlich weiter im Infraroten liegt, treten in der genannten Kanalkombination rote Farbtöne auf. Daher ist es naheliegend, zur Vegetationssegmentierung den roten Farbton heranzuziehen. Ausschließlich darauf aufbauende Verfahren werden in Stadtgebieten nicht immer zufriedenstellende Ergebnisse liefern können. Denn auch manche Hausdächer (Ziegel) weisen einen ähnlich roten Farbton auf. Doch wie zu Anfang des Kapitels erwähnt, kommt die Farbtonsegmentierung in Kombination mit der zuvor beschriebenen NDVI-Segmentierung zum Einsatz.

Ähnlich wie bei MAYER (1996) soll die Segmentierung aufgrund eindimensionaler Farbtonhistogramme erfolgen. Dazu muss aus dem RGB-Farbkomposit die Farbtonkomponente abgeleitet werden. In anderen Farbräumen als dem RGB-Farbraum repräsentiert der Farbton eine Dimensi-

on. So beruhen der HSV-, HSI- bzw. HSL-Farbraum (*Hue, Saturation, Value/Intensity/Luminance*) auf dem Farbton, der Farbsättigung sowie der Intensität/Luminanz (Definition des HSV-Farbraums durch SMITH (1978)). Diese Farbräume werden durch zylindrische Koordinaten aufgespannt. Der Farbton wird durch einen Winkel repräsentiert, während die Sättigung die Länge des Zylinderradius und die Intensität die Länge des Zylinders angibt (s. Abbildung 4.7). Die Überführung von RGB-Werten in den HSV-Farbraum erfolgt anhand Algorithmus 3, eine lineare Transformation existiert hierzu nicht (s. z. B. FOLEY (1990, S. 592 ff.), PITAS (2000, Kapitel 1.7)).

**Eingabe:** R, G, B

**Ausgabe:** H, S, V

RGB2HSV(R, G, B)

```
(1)  V ← MAX(R, G, B)
(2)  W ← MIN(R, G, B)
(3)  delta ← V - W
(4)  if V = 0
(5)    S ← 0
(6)    H ← -1
(7)    return H, S, V
(8)  S ← delta / V
(9)  if R = V
(10)   H ← 60(G - B)/delta
(11) else if G = V
(12)   H ← 120 + 60(B - R)/delta
(13) else
(14)   H ← 240 + 60(R - G)/delta
(15) if H < 0
(16)   H ← H + 360
(17) return H, S, V
```

**Algorithmus 3:** Farbraumkonvertierung von RGB zu HSV<sup>8</sup>.

Da die Vegetationssegmentierung ausschließlich aufgrund des Farbtons erfolgen soll, genügt es, diese Komponente anhand des RGB-Tripels zu berechnen. Die Farbtonwinkel werden in einem Histogramm aufgetragen. Nach Glättung des Histogramms durch Mittelwertfilterung wird das Maximum im 0°-Bereich bestimmt. Um dieses Maximum herum wird nun innerhalb vorgegebener Grenzen das Intervall bestimmt, in dem alle Histogrammeinträge über einer bestimmten Schwelle liegen (in der Implementierung werden 25 % des Maximums angesetzt). Daraus ergibt sich eine automatisch bestimmte Vorgabe zur Charakterisierung des roten Farbtonbereichs. Diese Grenzen dienen dem Nutzer als Orientierungshilfen und können (z. B. nach Begutachtung von Segmentierungsergebnissen) beliebig modifiziert werden.

Da bei dieser Segmentierung nur auf die Farbtonkomponente des HSV-Farbraums zurückgegriffen



wird, also die Intensität des Signals außer Acht gelassen wird, erzielt man damit eine gewisse Beleuchtungsinvarianz. Gerade bei radiometrisch nicht absolut kalibrierten Bilddaten ist dies von Bedeutung.

### 4.2.3 Kombination von NDVI und Farbtonsegmentierung zur Identifikation von Vegetation

Durch die spektrale Auslegung der HRSC-A auf geologische Fragestellungen bei der Marsbeobachtung ist, auch bei einer spektralen Simulation eines Rotkanals, keine zur HRSC-AX gleichwertige Trennbarkeit von Bäumen und Vegetation zu erwarten. Dies wird durch die Untersuchungen in Abschnitt 4.4.1 untermauert. Für die HRSC-A ist der NDVI nur bei „reinen“ Vegetationspixeln ausreichend hoch. Schatten, die großen Blickwinkelunterschiede sowie die daraus resultierenden größeren Verzerrungen bei der Orthokorrektur führen leicht zu einem vegetationsuntypischen NDVI-Wert. Ein Herabsetzen der NDVI-Schwelle resultiert jedoch in einer schlechten Differenzierung von Vegetation und Nicht-Vegetation.

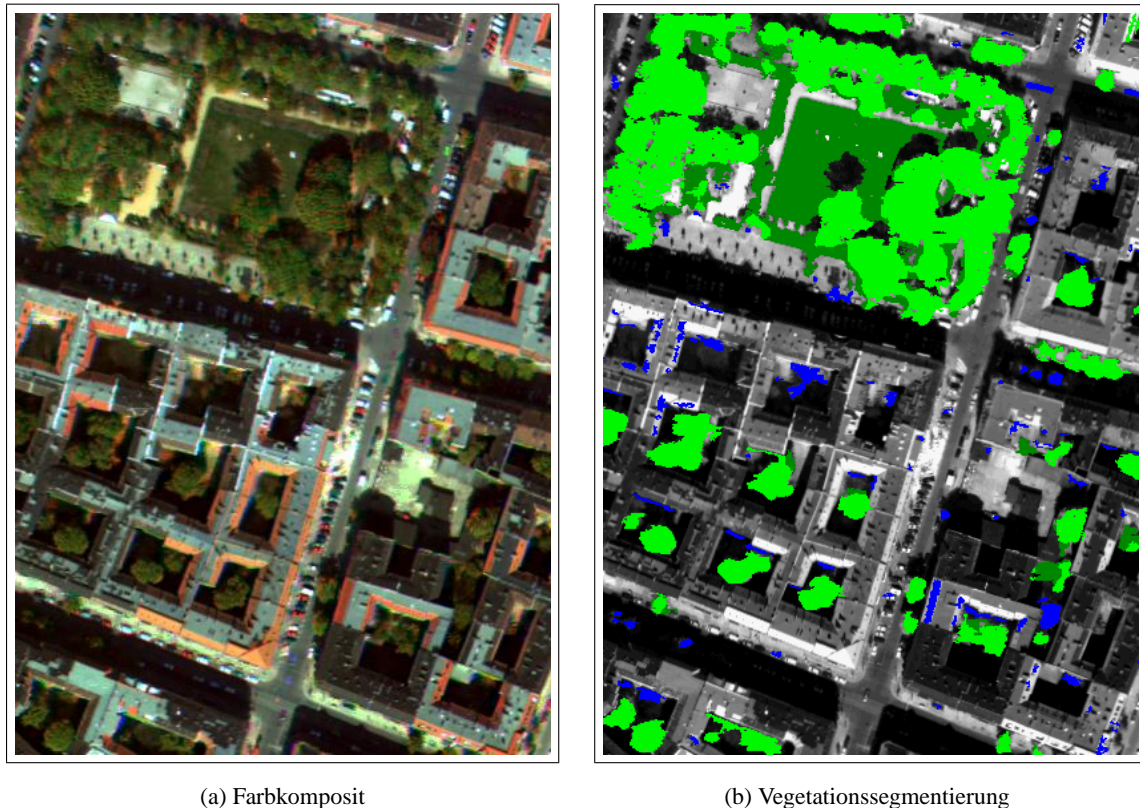
Bei einer Segmentierung des roten Farbtons sind die Segmente meist recht ausgedehnt und zusammenhängend. Allerdings weisen auch einige Hausdächer in der betrachteten Kanalkombination denselben roten Farbton auf wie Vegetation. Daher bietet sich eine Kombination der beiden Herangehensweisen an: Die Farbtonsegmentierung definiert die Segmente und durch den NDVI-Wert wird die Qualität der Segmente beurteilt. Algorithmus 4 gibt das Vorgehen zur Vegetationssegmentierung wieder.

Für jedes Rotsegment (in Algorithmus 4 bezeichnet als  $\text{seg}_i \in \mathcal{S}_{\text{rot}}$ ) die zugehörigen NDVI-Werte gemittelt. Ist der mittlere NDVI eines Segments größer als ein vorgegebener Schwellwert (Zeile 7), so wird es als Vegetation akzeptiert, andernfalls verworfen. Der Schwellwert  $\theta_{\text{NDVI}}$  ist frei wählbar; standardmäßig ist er in der aktuellen Implementierung auf 0,1 angesetzt.

**Eingabe:** IR, RE, BL, ND, RO (HRSC-A)  
**Ausgabe:**  $\mathcal{S}_{\text{veg}}$   
 VEGSEGM(IR, RE, BL, ND, RO)

- (1)  $\mathcal{S}_{\text{veg}} \leftarrow \emptyset$
- (2)  $\text{ROT}_{\text{sim}} \leftarrow \text{SIMULATERED}(\text{ND}, \text{RO})$
- (3)  $\text{NDVI} \leftarrow \text{CALCULATENDVI}(\text{ROT}_{\text{sim}}, \text{RO})$
- (4)  $H \leftarrow \text{HUE}(\text{IR}, \text{GR}, \text{BL})$
- (5)  $\mathcal{S}_{\text{rot}} \leftarrow \text{REDSEGM}(H, H_{\text{min}}, H_{\text{max}})$
- (6) **foreach**  $\text{seg}_i \in \mathcal{S}_{\text{rot}}$
- (7)     **if**  $\text{MEAN}(\text{NDVI}(\text{seg}_i)) > \theta_{\text{NDVI}}$
- (8)          $\mathcal{S}_{\text{veg}} \leftarrow \mathcal{S}_{\text{veg}} \cup \text{seg}_i$
- (9) **return**  $\mathcal{S}_{\text{veg}}$

**Algorithmus 4:** Berechnung einer Vegetationssegmentierung mit multispektralen HRSC-A Daten.



**Abbildung 4.8:** Das bereits in den letzten Abbildungen (vgl. Abb. 4.5(c)) verwendete HRSC-A Farbkomposit (a) und der zugehörige Nadirausschnitt überlagert mit dem Resultat der Vegetationssegmentierung (b). Die grün und blau dargestellten Flächen resultieren aus der Segmentierung des roten Farbtons. Durch die Integration des NDVI und einer Mindestgröße ergeben sich die endgültigen Vegetationssegmente (grün). Diese können anhand des nDOM wiederum differenziert werden in erhöhte (hellgrün) und niedere Vegetation (dunkelgrün).

Da die Farbtonsegmentierung auf anderen Spektralkanälen basiert als das NDVI-Verfahren, ergänzen sich die beiden Methoden auch in Bezug auf die Ausgangsdaten. Komplementäre multispektrale Information kann so ideal genutzt werden. Abbildung 4.8 zeigt einen Bildausschnitt zusammen mit dem Ergebnis der Vegetationssegmentierung. Dabei sind zum einen die allein durch den Farbton segmentierten Flächen dargestellt, zum anderen die Untermenge davon, die auch den NDVI-Test besteht. Wird das normierte Höhenmodell einbezogen, kann mittels Schwellwert die erhöhte Vegetation leicht von der niederen getrennt werden. In der Abbildung sind diese beiden Vegetationskategorien durch zwei unterschiedliche Grüntöne dargestellt.

Das in der Farbtonsegmentierung existierende Problem der Fehlzuzuordnung einiger Hausdächer wird durch die Kombination der beiden Verfahren stark unterdrückt. Jedoch verbleiben auch hier vereinzelt als Vegetation eingeordnete Dachsegmente (s. Abb. 4.8(b)). Eine abschließende Bewertung

der sowohl kombiniert als auch separat durchgeführten Ansätze zur Vegetations-/Gebäudesegmentierung erfolgt in Abschnitt 4.4.

#### 4.2.4 Vereinfachtes Vorgehen für HRSC-AX Daten

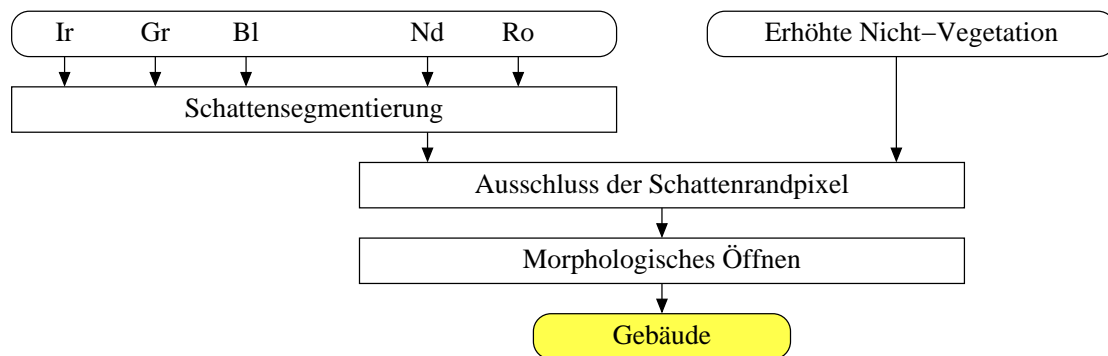
Die Auslegung der Bandbreiten der HRSC-AX Kanäle entspricht den Anforderungen der Erdfernerkundung (s. Kapitel 2). Wie Abbildung 4.3 zu entnehmen ist, flankieren der rote und der infrarote HRSC-AX Kanal den für die NDVI-Berechnung relevanten Bereich um 700 nm. Sie können so direkt zur Bestimmung des NDVI herangezogen werden. Der Blickwinkelunterschied dieser beiden Kanäle beträgt lediglich  $2,3^\circ$ . Blickwinkelbedingte Effekte haben deswegen einen wesentlich geringeren Einfluss als bei der HRSC-A; dort sind der panchromatische Nadir sowie der Rotkanal am NDVI beteiligt, was zu einem Blickwinkelunterschied von  $15,9^\circ$  führt<sup>9</sup>. Je größer die Blickwinkeldifferenz zur Orthoperspektive ist, desto stärker sind die Kanäle von der Orthokorrektur (s. Abschnitt 2.3.5) betroffen. Die damit verbundenen Verzerrungen wirken sich auch auf den NDVI aus — gerade bei Objekten mit einer ausgeprägten Oberflächenmorphologie wie Bäumen. Zusammenfassend ist bei der HRSC-AX eine bessere Trennbarkeit von Bäumen und Gebäuden durch das NDVI-Kriterium zu erwarten. Die in Abschnitt 4.4.1 folgenden Untersuchungen diesbezüglich entsprechen dieser Vermutung.

Wie bereits in den vorhergehenden Abschnitten erwähnt, wird ein zweites Kriterium zur flächenhaften Erfassung von Bäumen mit HRSC-A Daten benötigt. Eine ausschließlich aufgrund des NDVI-Kriteriums durchgeführte Segmentierung, die Baumflächen gut erfasst, würde zu einer zu hohen Fehlalarmrate führen. Einerseits ist die unzufriedenstellende Erfassung von Bäumen bei alleiniger Verwendung des NDVI durch die hohe geometrische Auflösung bedingt. Neben Blattpixeln sind auch Ast- und Bodenpixel sichtbar. Von Effekten dieser Art sind Daten der HRSC-AX jedoch in gleichem Maß betroffen. Andererseits kann der NDVI von HRSC-A Daten aufgrund der angeführten Kriterien die vegetationstypischen spektralen Charakteristika weniger gut erfassen als ein auf HRSC-AX Daten aufbauender Index.

Mehrere Experimente haben gezeigt, dass bei Verwendung von Daten der HRSC-AX alleine das NDVI-Kriterium zur Vegetationssegmentierung genügt. Die Hinzunahme des Farbtonkriteriums ist möglich, führt aber zu keinen nennenswerten Verbesserungen bezüglich der Form der Baumsegmente. Der in Abbildung 4.2 gezeigte Ablauf zur Vegetationssegmentierung kann bei HRSC-AX Daten also deutlich vereinfacht werden.

---

<sup>9</sup>Statt des Nadirs ließe sich auch der nach vorne blickende panchromatische Photometriekanal verwenden. Ein Blickwinkelunterschied von  $2,9^\circ$  wäre die Folge. Die Photometriekanäle werden jedoch standardmäßig nur zur Bildzuordnung verwendet und nicht orthokorrigiert.



**Abbildung 4.9:** Schematische Darstellung zur Nachverarbeitung der Gebäudesegmente. Eingangsdaten sind die Multispektralkanäle sowie die in Abschnitt 4.2 erarbeitete Segmentierung in erhöhte Nicht-Vegetationspixel.

### 4.3 Nachverarbeitung der Gebäudesegmente

Betrachtet man die durch die vorhergehenden Schritte erzielten Segmentierungsergebnisse näher (s. Abb. 4.11), so fällt auf, dass die Form der Regionen oft durch Schatten beeinflusst wird. Gerade die besonders interessierenden erhöhten Regionen weisen — entsprechend dem Sonnenstand und der Helligkeit — an ihren Grenzen einen Schattenwurf auf. Der Schatten selbst befindet sich in Realität meist auf dem Grund. Wegen der in Abschnitt 2.3.4 und 2.5.2 erwähnten Interpolationsfehler im HRSC-Oberflächenmodell kann Schatten, der durch Höhengsprünge verursacht wurde, jedoch auch den erhöhten Regionen zugeordnet werden.

Die Vegetationssegmente enthalten weniger störende Schattenränder, da Schatten meist nicht den Segmentierungskriterien für Vegetation genügt. Gebäudesegmente ergeben sich jedoch aus den verbleibenden Nicht-Vegetationspixeln der erhöhten Regionen. Dadurch werden die in der Vegetationssegmentierung abgewiesenen Schattenpixel automatisch der Gebäudekategorie zugeordnet. Um eine genauere Gebäudesegmentierung zu erzielen, ist deshalb die Detektion der Schattenregionen erforderlich. Auf diese Thematik wird im folgenden Abschnitt eingegangen. Eine abschließende Nachverarbeitung der schattenreduzierten Gebäudesegmente wird dann in Abschnitt 4.3.2 vorgestellt. Abbildung 4.9 gibt einen schematischen Überblick über die Behandlung der aus den vorangegangenen Abschnitten resultierenden Segmente erhöhter Nicht-Vegetationspixel.

#### 4.3.1 Schattensegmentierung

In der Bildverarbeitung wird die Detektion von Schatten hauptsächlich zu zwei Zwecken betrieben. Zum einen, um eine bessere Separierung der Objektklassen zu ermöglichen; zum anderen, um aus der Form oder Position des Schattens Rückschlüsse auf das schattenwerfende Objekt („Shape from Shading“)<sup>10</sup>, bzw. die Beleuchtungsquellen („Inverse Lighting“, MARSCHNER UND GREENBERG

<sup>10</sup>Diskussion der gebräuchlichen Algorithmen zu diesem Thema bei HORN UND BROOKS (1989).

(1997), und „Radiometric Reconstruction“, LUONG ET AL. (2002)) zu ziehen. In dieser Arbeit dient die Schattendetektion dem erstgenannten Ziel, der Verfeinerung der Gebäudesegmentierung. Die in der Literatur dazu existierenden Algorithmen<sup>11</sup> lassen sich nach PRATI ET AL. (2001) danach untergliedern, ob das Vorgehen statistisch (parametrisch/nicht parametrisch) oder deterministisch (modellbasiert/modellfrei) ist. Ferner kann nach der Wahl der verwendeten Merkmale unterschieden werden. Spektrale Charakteristika wie Grauwert oder Farbe, räumliche Merkmale (Zusammenhänge in einer lokalen Nachbarschaft oder regionenbasiert) sowie bei Filmsequenzen Merkmale, die zeitliche Zusammenhänge erfassen. Diese Verfahren zur Schattendetektion lassen sich jedoch nicht ohne Weiteres auf das vorliegende Szenario übertragen.

So wird bei BARNARD UND FINLAYSON (2000), MIKIC ET AL. (2000) von einem diagonalen Kameramodell ausgegangen. D. h. alle existierenden Farbwerte sollten sich durch Multiplikation mit einer für die Kamera charakteristischen Diagonalmatrix ineinander überführen lassen. Ein Hinweis auf das Vorhandensein eines Schattensegments neben einem beleuchteten Segment ist eine dem Diagonalmodell entsprechende Abschwächung der Farbwerte (BARNARD UND FINLAYSON 2000).

Ähnlich dem Diagonalmodell ist auch die grundlegende Annahme bei FUNKA-LEA UND BAJC-SY (1995). Die Schattenpenumbra (Halbschatten, erzeugt durch Umgebungslicht) einer Farbe wird in einem linearen Farbraum wie dem RGB abgebildet auf ein lineares Cluster. Die Pixel in einer Nachbarschaft um zuvor bereits segmentierte (bewegte) Objekte werden auf diese Eigenschaft hin überprüft. Als weiteres Kriterium dient in diesem Ansatz die Gestalt der Schattenfläche. Nur nahezu planare Flächen (ermittelt über Laser-Messungen) kommen als Schatten in Betracht. Voraussetzung dafür ist jedoch, dass die Fläche, auf die der Schatten projiziert wird, planar ist.

Bei der Erkennung bewegter Objekte in Videosequenzen dient die Detektion von Schatten der Trennung von Objekt und zugehörigem direkt benachbarten Schatten. Typischerweise geht der Schattendetektion in diesem Fall eine Segmentierung der bewegten Objekte voraus, was die Komplexität der Detektionsproblematik erheblich reduziert. Ferner stehen meist Referenzaufnahmen ohne die Beschattung durch bewegte Objekte zur Verfügung (ROSIN UND ELLIS 1995, HORPRASERT ET AL. 2000, PRATI ET AL. 2001, s. z. B.). Die Abschattung einer Objektfarbe wird in den zitierten Ansätzen über die niedrigere Luminanz im Vergleich zum Referenzbild erkannt. Bei HORPRASERT ET AL. (2000) und PRATI ET AL. (2001) wird zusätzlich angenommen, dass der Farbwert dabei nahezu unverändert bleibt. Bei ROSIN UND ELLIS (1995) müssen die Schattenregionen zusätzlich statistische Kriterien in Bezug auf ihre Grauwerte erfüllen.

In dieser Arbeit erfolgt die Schattendetektion durch einen deterministischen, modellfreien Ansatz auf Basis eines einfachen radiometrischen Kriteriums, die geringe Reflektanz im Vergleich zur Umgebung. Beim Einsatz von Multispektraldaten sind alle Kanäle gleichermaßen davon betroffen. Meist sind die Schattenregionen sogar die dunkelsten Regionen der Aufnahme. Diese simplen radiometrischen Schattencharakteristika werden in dem Helligkeitsmaß  $V$  erfasst („Value“, s. auch Anmerkun-

---

<sup>11</sup>FUNKA-LEA (1994) diskutiert zahlreiche Algorithmen zur Schattendetektion.

gen zum HSV-Farbraum in Abschnitt 4.2.2 und Algorithmus 3 zur Farbraumkonvertierung von RGB zu HSV). Die Helligkeit ist für alle Pixel definiert als Maximum über alle  $M$  Multispektralkanäle (s. SMITH 1978):

$$V(x, y) = \max_i f_i(x, y), \quad i = 1, \dots, M \quad (4.5)$$

Bei dieser Helligkeitsdefinition werden alle Spektralkanäle gleich gewichtet. Eine Gewichtung entsprechend der spektralen Bandbreite der Kanäle ist ebenfalls denkbar.

Anhand eines geeigneten Helligkeitsschwellwerts werden somit die in allen Kanälen dunklen Pixel selektiert<sup>12</sup>. Abbildung 4.10 zeigt einen Bildausschnitt und das Resultat der Schattensegmentierung. Offensichtlich ist durch Anwendung des Schwellwertes nicht ausgeschlossen, dass auch andere Objekte mit geringer Reflektanz erfasst werden. JIANG UND WARD (1992) legen deshalb die Schwelle relativ zu einer Grundhelligkeit fest, die sich zeilenweise aus einer Helligkeitsgeraden durch die ersten und letzten Pixel ergibt. Die Erfassung der Helligkeitsvariation ist somit stark von den die Helligkeitsgerade bestimmenden Objekten abhängig.

GWINNER UND SCHAALE (1997) ermitteln den optimalen Schwellwert zur Trennung von Schatten und Nicht-Schatten bei multispektralen Daten. Dazu wird für jeden Kanal durch Minimierung der Fehlklassifikationsrate anhand von Testdaten der beste Schwellwert bestimmt. Da die daraus resultierenden Schattensegmente meist zu klein sind (die Penumbra wird nicht erfasst), erfolgt eine Nachverarbeitung. Basierend auf der Gleichförmigkeit der Gradientenrichtung in allen Spektralkanälen werden benachbarte Pixel durch Regionenwachstum in das jeweilige Segment einbezogen (GWINNER 1996).

Wie zu Anfang des Abschnitts beschrieben, ist das Ziel nicht eine verlässliche Schattendetektion, sondern lediglich das Erkennen von Schatten an den Grenzen der Gebäudesegmente („Ausschluss der Schattenrandpixel“ in Abbildung 4.9). Aus diesem Grund werden, ähnlich zu IRVIN UND MCKEOWN (1989) und LIOW UND PAVLIDIS (1990), die dunklen Bildbereiche an die Existenz angrenzender Gebäuderegionen gekoppelt. Durch diese Reduktion des relevanten Bildbereichs wird auch die Möglichkeit für Schattenfehldetektionen erheblich eingeschränkt. Gebäudesegmentgrenzen und die aus der Helligkeitsschwelle resultierenden Schattensegmente werden dazu logisch UND-verknüpft. Die zu eliminierenden Schattensegmente sind dadurch gekennzeichnet, dass sie Randpixel der „erhöhten Nicht-Vegetationssegmente“ enthalten müssen. Beschattungen auf der Dachfläche, z. B. durch Kamine oder Antennen, haben bei diesem Vorgehen also keine Auswirkung. Lediglich Schatten, die durch noch höhere benachbarte Gebäude hervorgerufen werden, können nicht sinnvoll abgefangen werden. In Abbildung 4.11 ist das Resultat der Schattenelimination dargestellt. Die in Abbildung 4.11(b) cyan markierten Regionen sind Schatten am Rand der Gebäudesegmente. Auf eine weitere Bereinigung der so überarbeiteten Gebäudesegmente geht der folgende Abschnitt ein.

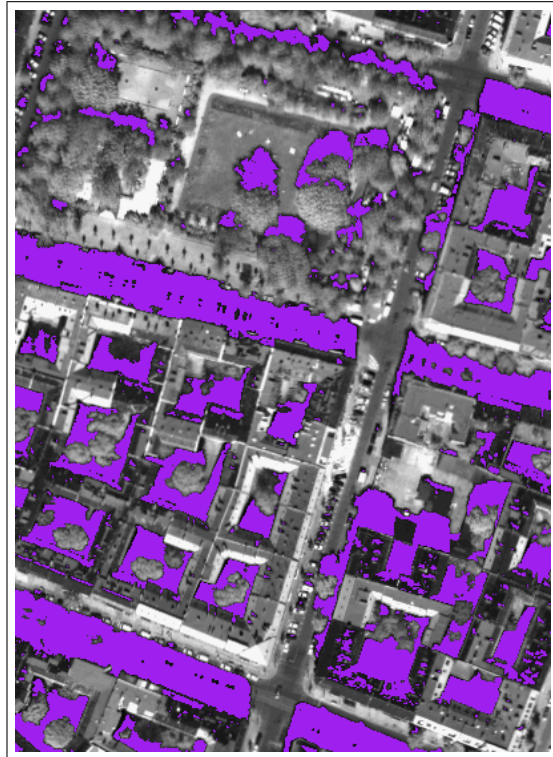
---

<sup>12</sup>In der Implementierung erfolgt dies durch rein manuelle Angabe eines Schwellwerts. Zur Automatisierung könnte die Schwellwertbestimmung aber auch basierend auf dem Helligkeitshistogramm analog zu dem in Abschnitt 3.3 beschriebenen Vorgehen erfolgen.





(a) Farbkomposit



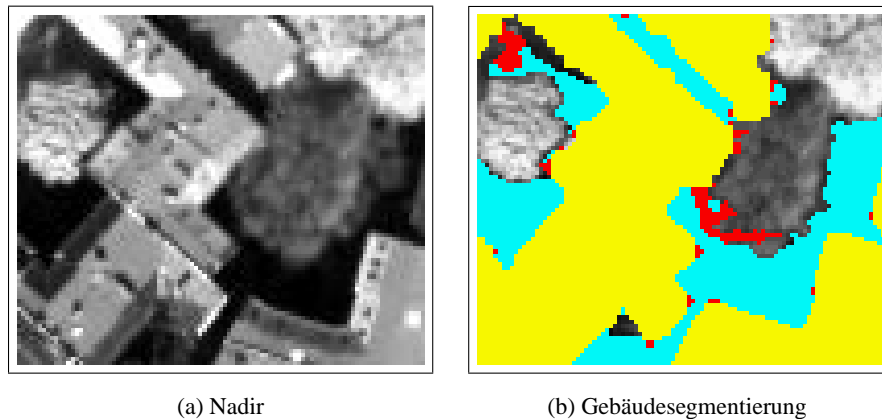
(b) Schattensegmentierung

**Abbildung 4.10:** Das bereits in den letzten Abbildungen verwendete HRSC-A Farbkomposit (a), s. auch Abb. 4.5(c), und der zugehörige Nadirausschnitt violett überlagert mit dem Resultat der Schattensegmentierung (b).

#### 4.3.2 Morphologisches Öffnen

Gebäudesegmente werden in den vorangegangenen Abschnitten nicht konstruktiv bestimmt (in dem Sinn, dass die Segmente aufgrund charakteristischer Gebäudeeigenschaften abgeleitet werden), sondern durch den Ausschluss negativer Eigenschaften. Dazu werden zuerst Vegetationssegmente bestimmt. Die initialen Gebäudesegmente ergeben sich dann aus den erhöhten Segmenten ohne Vegetation. Anschließend werden die Schattenbereiche detektiert und ebenfalls von den Gebäudesegmenten subtrahiert. Bei diesem „Beschneiden“ der Segmente erhöhter Pixel kann es durchaus passieren, dass in den verbleibenden Gebäudesegmenten an der Konturlinie Artefakte in Form von Ausläufern zurückbleiben. In Abbildung 4.11 ist die Veränderung des Gebäudesegmente durch Ausschluss der Schattenpixel exemplarisch dargestellt (alle Farben: initiale Gebäudesegmente, cyan: Schattenbereiche an den Segmenträndern).

Zur Beseitigung der Artefakte werden die Gebäudesegmente durch morphologisches Öffnen nachbearbeitet. Die Segmente werden mit einem kreisförmigen strukturierenden Element (in der Implementierung mit einem Durchmesser von 5 Pixeln) morphologisch erodiert, um die unerwünschten



**Abbildung 4.11:** Bildbeispiel zur morphologischen Nachverarbeitung der Gebäudesegmente. Links (a) ein Ausschnitt aus dem Nadirkanal und rechts (b) zusätzlich die farbliche Überlagerung der Gebäudesegmente. Die Vereinigung aller farblich dargestellten Flächen repräsentiert erhöhte Segmente ohne Vegetation. Cyan dargestellt sind die Randteile dieser vorläufigen Gebäudesegmente, die als Schatten identifiziert werden. Die von der morphologischen Nachbehandlung betroffenen Pixel sind rot dargestellt. Resultat dieser Überarbeitungsschritte sind die gelb dargestellten Gebäuderegionen.

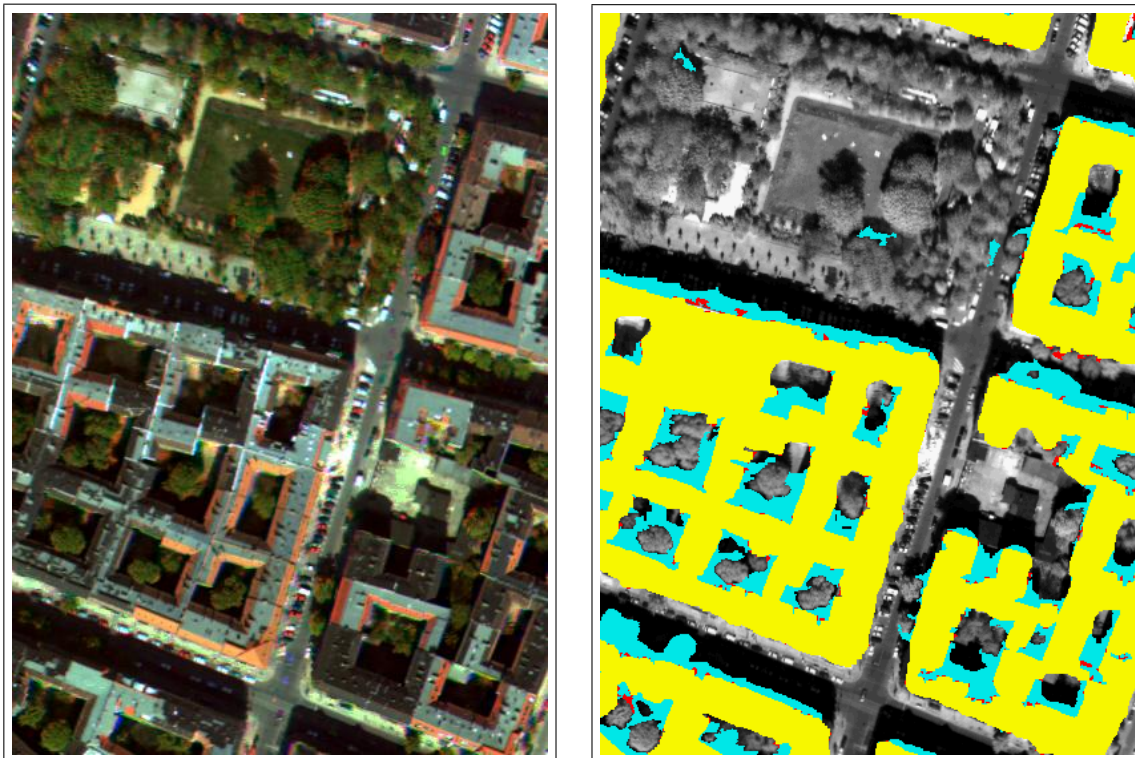
Auswüchse vom Hauptsegment abzuspalten. Dann werden Segmente, die kleinere Abmessungen als ein (erodiertes) Haus haben, eliminiert. Um wieder die ursprüngliche Form der Gebäudesegmente (ohne die nun gelöschten Ausläufer) zu erhalten, werden die Segmente abschließend mit dem gleichen strukturierenden Element dilatiert. Die erfolgreich eliminierten Segmentausläufer sind in Abbildung 4.11(b) rot dargestellt. Das resultierende endgültige Gebäudesegment ist gelb gekennzeichnet. In Abbildung 4.12 ist dasselbe Vorgehen zur Gebäudesegmentierung an einem größeren Bildausschnitt illustriert.

## 4.4 Evaluation

Abbildung 4.2 gibt einen Eindruck von dem Resultat der Trennung von Bäumen und Häusern durch die in diesem Kapitel vorgestellten Methoden. Bis auf wenige Ausnahmen ist die Unterteilung erfolgreich. In Abschnitt 4.4.1 wird allgemein die Trennbarkeit von Bäumen und Gebäuden durch Berechnung des minimalen Fehlers bei Klassifikation der eingeführten Merkmale bewertet. Angaben zur Genauigkeit der vorgestellten Baum- und Gebäudedetektion folgen in den Abschnitten 4.4.2 und 4.4.3.

Zur Ermittlung der Genauigkeit wird die Handsegmentierung eines Berliner Stadtgebiets als Referenz herangezogen. Es umfasst mehrere Wohnblocks, Straßen mit Baumbegrünung und eine kleine Grünanlage (s. Abbildung 4.2). Die manuelle Auswertung basiert auf denselben Bilddaten, die auch in der automatisierten Vorgehensweise zum Einsatz kommen: Fünf orthokorrigierte Spektralkanäle





(a) Farbkomposit

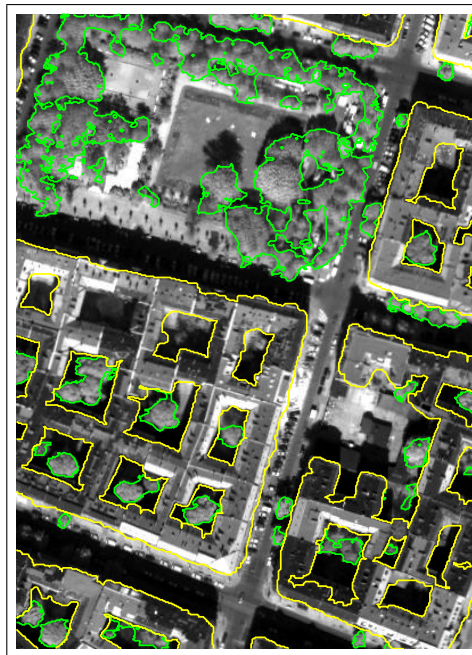
(b) Gebäudesegmentierung

**Abbildung 4.12:** Das bereits in den letzten Abbildungen verwendete HRSC-A Farbkomposit (a) und der zugehörige Nadirausschnitt überlagert mit dem Resultat der Gebäudesegmentierung (b). Die Farben entsprechen der Darstellung in Abbildung 4.11.

der HRSC-A sowie das DOM. Der Bildausschnitt misst  $1170 \times 970$  Pixel bei 0,5 m Pixelauflösung. Die etwa eine Million Pixel entsprechen somit einer Grundfläche von  $0,25 \text{ km}^2$ . In der Handsegmentierung werden die beiden Klassen „Gebäude“ und „Bäume“ erfasst, also die Resultate der automatisierten Segmentierung erhöhter Regionen.

Auf eine vollständige Auswertung mit Daten der HRSC-AX wird verzichtet. Jedoch folgt in Abschnitt 4.4.1 eine Diskussion der Separierbarkeit von Gebäuden und Vegetation durch einen NDVI-basierten Ansatz.

Die Laufzeiten der in diesem Kapitel besprochenen Routinen fallen gegenüber der manuellen Interaktionszeit (z. B. Auswahl der entsprechenden Dateien, Parameterwahl) und gegenüber den in den Kapiteln 3, 5 und 6 behandelten Ansätzen kaum ins Gewicht. Aus diesem Grund wird auf eine Messung der Programmlaufzeiten in diesem Kapitel verzichtet. Für eine praktische Umsetzung ist eher die gesamte Bearbeitungszeit im Vergleich zu einer rein manuellen Vorgehensweise von Interesse. Zur Vektorisierung der Gebäude und Bäume in dem erwähnten Testdatensatz hat ein in der



**Abbildung 4.13:** Resultat der Gebäude- und Vegetationssegmentierung. Der Bildausschnitt ist überlagert mit den Konturen von erhöhter Vegetation und Gebäuden, die sich aus den in Abbildung 4.8(b) und 4.12(b) dargestellten Segmenten ergeben.

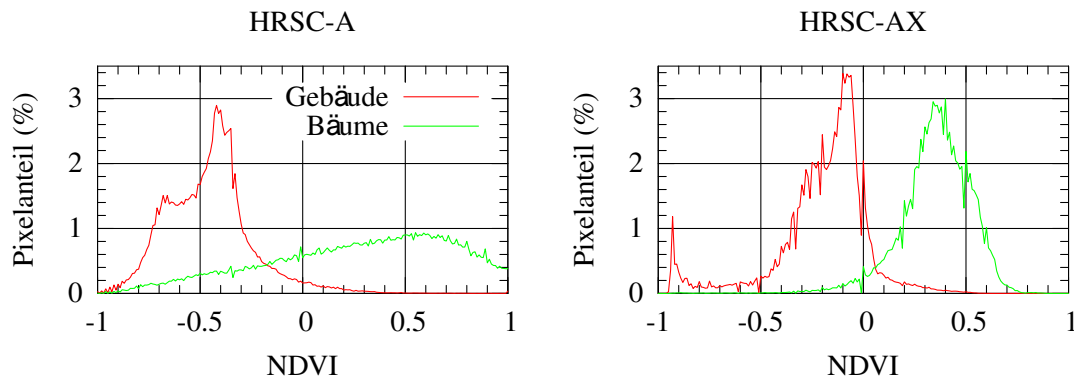
Bildinterpretation unerfahrener Laie etwa 20 h benötigt<sup>13</sup>. Eine Abarbeitung der in diesem Kapitel vorgeschlagenen Routinen ist auch für unerfahrene Nutzer innerhalb einer Stunde zu erledigen<sup>14</sup>. Bei beiden Vorgehensweisen ist die Bearbeitungszeit durch Erfahrung und Routine deutlich senkbar. Der automatisierte Ansatz baut jedoch im Gegensatz zum manuellen Vorgehen auf einer vorhandenen Segmentierung der erhöhten Objekte auf. Somit sind im Fall von unebenem Gelände die in Abschnitt 3.4.1 angegebenen Laufzeiten zur DOM-Normierung zu berücksichtigen.

#### 4.4.1 Trennbarkeit von Bäumen und Gebäuden bei HRSC-A und HRSC-AX

Für die folgenden Betrachtungen kommen neben dem vorab erwähnten HRSC-A Datensatz auch HRSC-AX Daten eines Münchener Stadtgebiets zum Einsatz. Beide Datensätze sind auf 0,5 m Pixelauflösung reduziert. Ferner existiert eine Handsegmentierung für Gebäude- wie auch für Baumregionen zu beiden Datensätzen. Aussagen zur Trennbarkeit von Vegetation und Gebäuden beziehen sich somit nur auf erhöhte Regionen. Bei den HRSC-A Daten wird zur NDVI-Berechnung ein künst-

<sup>13</sup>Durch das manuelle Vektorisieren der Objekte liegen für die Gebäudesegmente gleichzeitig geradlinige Konturlinien vor. Um mit dem propagierten Vorgehen vergleichbare Ergebnisse zu erhalten, muss noch eine Anpassung der Gebäudedekonturen an Bildkanten gemäß Kapitel 6 erfolgen. Diese Bearbeitungszeiten wären also ebenfalls zu berücksichtigen.

<sup>14</sup>Der bearbeitete Bildausschnitt ist größer ( $2\,000 \times 2\,000$  Pixel) als das davon zur Auswertung herangezogene Teilgebiet. Da die Programmlaufzeiten gegenüber der manuellen Interaktion jedoch vernachlässigbar sind, spielt dies hier keine Rolle.



**Abbildung 4.14:** Relative Pixelhäufigkeiten in Abhängigkeit vom NDVI bei HRSC-A und HRSC-AX Daten.

licher Rotkanal berechnet, bei Daten der HRSC-AX kommen der originale Rot- und Infrarotkanal zum Einsatz.

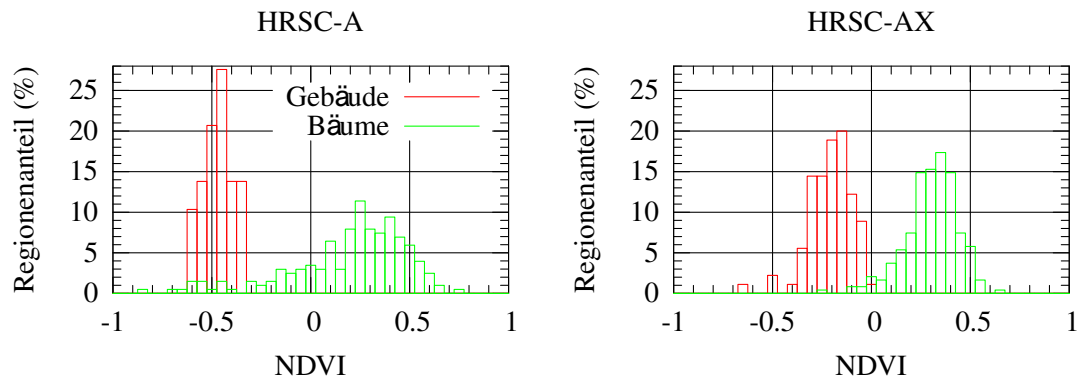
Neben den unterschiedlichen (doch inhaltlich ähnlichen) Zielgebieten und den verschiedenen Beleuchtungsverhältnissen ist vor allem auf den unterschiedlichen Entwicklungsstand der Vegetation hinzuweisen. Die beiden Datensätze wurden zwar jeweils Anfang April (einmal im Jahr 2000, das andere Mal im Jahr 2002) aufgenommen, doch viele der Bäume im HRSC-AX München-Datensatz sind schwächer belaubt als die des Berlin-Datensatzes. Daher kann die nun folgende Gegenüberstellung nur als grober Anhaltspunkt gesehen werden.

Aus der Handsegmentierung der beiden Datensätze resultiert die in Abbildung 4.14 wiedergegebene pixelweise Verteilung der NDVI-Werte. Die Verteilung ist für die beiden zu trennenden Klassen separat aufgetragen. Die Häufigkeiten sind anteilig zur Gesamtzahl der Pixel einer Klasse.

Wie zu erwarten, ist die Verteilung der NDVI-Werte bei Vegetationspixeln für die beiden Sensoren sehr unterschiedlich. Bei den HRSC-A Daten weisen sie eine wesentlich höhere Streuung auf; die Vegetationspixel sind fast über den gesamten NDVI-Wertebereich verteilt. Eine Trennung der beiden Klassen ausschließlich aufgrund des NDVI-Kriteriums ist so wenig erfolgversprechend. Um dies zu quantifizieren, wird der minimale Klassifikationsfehler anhand der Handsegmentierung abgeschätzt. Der minimale Fehler ergibt sich aus der Bayes'schen Entscheidungsregel, die Klasse mit der größten A-posteriori-Wahrscheinlichkeit zu wählen (s. z. B. DUDA ET AL. 2001, Kapitel 2.1). Die A-posteriori-Wahrscheinlichkeit ist definiert als

$$p(\Omega_i|\mathbf{m}) = \frac{p(\mathbf{m}|\Omega_i)p_i}{\sum_j p(\mathbf{m}|\Omega_j)p_j}, \quad (4.6)$$

mit  $p(\mathbf{m}|\Omega_i)$  als der bedingten Wahrscheinlichkeit des Merkmals  $\mathbf{m}$  bei Vorliegen der Klasse  $\Omega_i$  und  $p_i$  als A-priori-Wahrscheinlichkeit der Klasse  $\Omega_i$ . Da durch die Handsegmentierung jedem Pixel eine Klasse zugewiesen ist, ergeben sich anhand der relativen Häufigkeiten Schätzungen für  $p(\mathbf{m}|\Omega_i)$  und  $p_i$  und somit auch für die A-posteriori-Verteilung.



**Abbildung 4.15:** Relative Häufigkeiten von Regionen in Abhängigkeit vom zugeordneten mittleren NDVI-Wert bei HRSC-A und HRSC-AX Daten.

Die pixelweise bestimmte Fehlerrate ist bei den HRSC-A Daten mehr als doppelt so hoch wie bei HRSC-AX Daten — trotz besserem Entwicklungsstand der in den HRSC-A Daten erfassten Vegetation. Die Fehlerraten sind in Tabelle 4.1 zusammengefasst<sup>15</sup>.

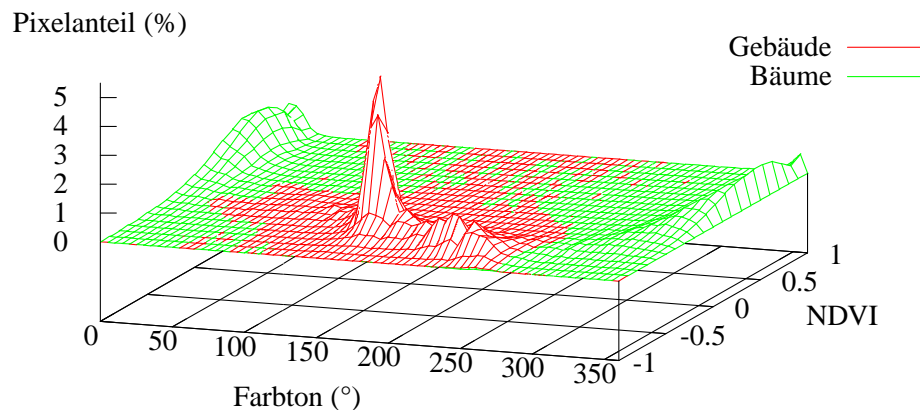
Klassifikation	HRSC-A (%)	HRSC-AX (%)
NDVI, pixelweise	8,70	3,82
NDVI-Mittel, regionenweise	4,76	1,24
NDVI und Farbton, pixelweise	4,92	—

**Tabelle 4.1:** Abschätzung des minimalen Klassifikationsfehlers bei der Trennung von Bäumen und Gebäuden für Daten der HRSC-A und HRSC-AX. Grundlage sind die Handsegmentierungen der jeweiligen Datensätze (die relativen Häufigkeiten für das jeweils verwendete Klassifikationsmerkmal sind in den Abbildungen 4.14, 4.15 und 4.16 wiedergegeben).

In Abschnitt 4.2.3 wird vorgeschlagen, den NDVI regionenweise zu bestimmen. Von jedem Segment wird der mittlere NDVI-Wert ermittelt. Die regionenweise Verteilung der NDVI-Werte in den Testdatensätzen wird in Abbildung 4.15 wiedergegeben. Die Zuordnung als Gebäude oder Vegetation erfolgt dann aufgrund des NDVI-Mittelwerts. Diese Mittelung sorgt für eine größere Robustheit gegenüber Ausreißerpixeln. Eine Senkung des minimalen Fehlers bei dem Zwei-Klassen-Problem ist die Folge. Bei den in Tabelle 4.1 angegebenen Fehlerraten ist anzumerken, dass die Anzahl der Regionen (gegenüber der pixelweisen Behandlung) für statistische Aussagen gering ist (Bäume:  $\leq 242$ , Gebäude:  $\leq 90$ ).

Eine zusätzliche Verbesserung ist durch die Hinzunahme eines weiteren Merkmals zu erwarten. In Abschnitt 4.2.2 wird für Daten der HRSC-A der Farbton der Infrarot-, Grün-, Blaukanalkombina-

<sup>15</sup>Die Klassifikationsentscheidung wird jeweils unter Einbeziehung der A-priori-Wahrscheinlichkeit der beiden Klassen getroffen. In den Darstellungen in Abb. 4.14–4.16 ist die A-priori-Wahrscheinlichkeit durch die Normierung nicht erfasst. Dadurch kann der Überlappungsbereich größer erscheinen als er bei der Bayes-Entscheidung ist.



**Abbildung 4.16:** Relative Pixelhäufigkeiten in Abhängigkeit vom Farbton und NDVI bei HRSC-A-Daten.

tion gewählt. Die relative Pixelverteilung des HRSC-A Datensatzes im zweidimensionalen Farbton-NDVI-Raum ist in Abbildung 4.16 dargestellt. Neben dem NDVI unterscheidet sich nun auch der mittlere Gebäudefarbton von etwa 170° (cyan-grau der Dachpappe) vom Rotton der Vegetation. Die Fehlerraten für den Bayes-Klassifikator sinken gegenüber dem eindimensionalen Vorgehen auf fast die Hälfte.

Eine regionenbasierte Auswertung macht aufgrund des geringen Stichprobenumfangs keinen Sinn. Zur Ermittlung der relativen Häufigkeiten müssten die wenigen Regionen nun auf ein zweidimensionales Histogramm verteilt werden.

Zusammenfassend ergibt sich folgendes Bild:

- Die HRSC-AX ist aufgrund der geänderten Multispektralkanäle zur Unterscheidung von Bäumen und Gebäuden gut geeignet. Mit einem NDVI-basierten Ansatz können die beiden Klassen erfolgreich getrennt werden.
- Mit einem regionenbasierten Vorgehen wird die Fehlerrate bei der Unterteilung von Bäumen und Gebäuden gesenkt.
- Die zusätzliche Nutzung des Farbtons bei Daten der HRSC-A trägt ebenfalls zur Senkung der Fehlerrate bei.

Die in diesem Abschnitt vorgestellte Gegenüberstellung der beiden HRSC-Generationen zur Unterscheidung von Gebäuden und Vegetation unterstreicht die Bedeutung der Auslegung der Multispektralkanäle. Bei der nun folgenden Auswertung der Baum- und Gebäudedetektion kommen ausschließlich Daten der HRSC-A zum Einsatz.

### 4.4.2 Baumdetektion

Zur Baumdetektion wird entsprechend dem in diesem Kapitel herausgearbeiteten Vorgehen verfahren: Basierend auf dem roten Farbton und dem NDVI-Wert werden die Vegetationssegmente bestimmt. Dies erfolgt einerseits kombiniert, wie in Abschnitt 4.2.3 vorgeschlagen, andererseits zum Vergleich nur mit jeweils einem der Verfahren. Aus den resultierenden Vegetationssegmenten werden die Segmente erhöhter Vegetation durch Verschneidung mit den aus dem nDOM ermittelten erhöhten Regionen bestimmt. Nach einer Entfernung zu kleiner Segmente ergeben sich dann die endgültigen Baumsegmente.

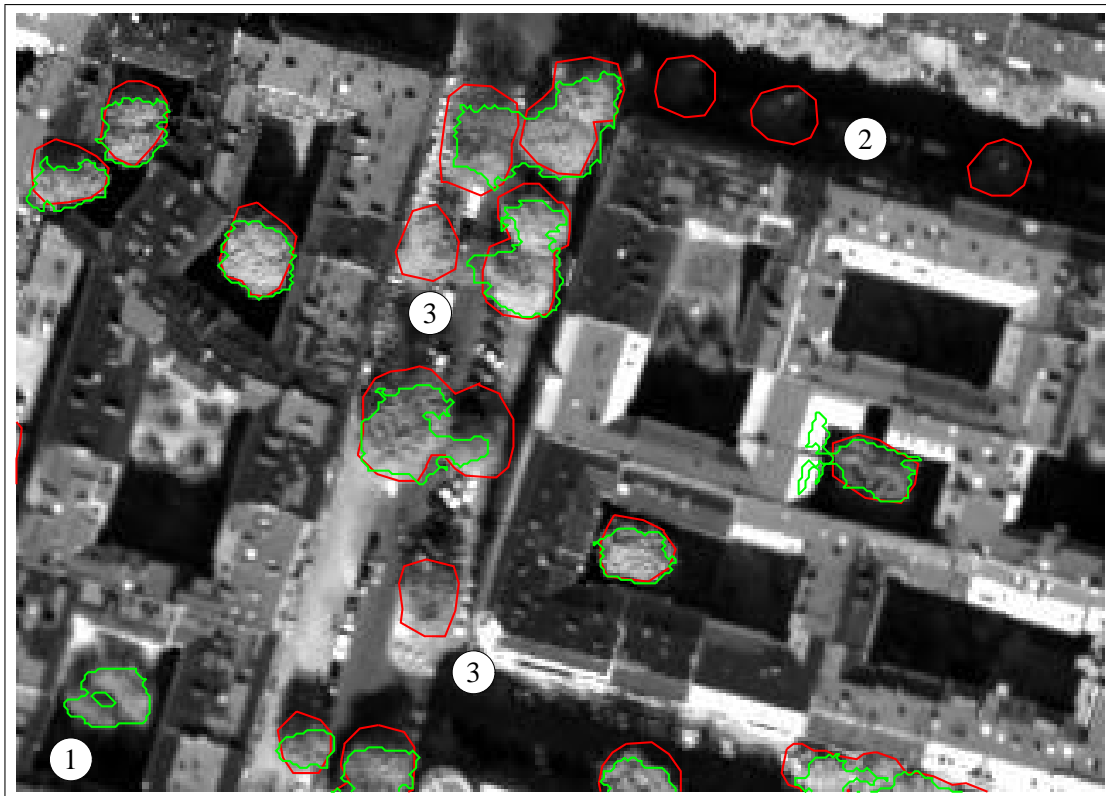
Während die Konturlinie eines Gebäudesegments (geradlinig, meist guter Kontrast zur Umgebung) manuell in Pixelgenauigkeit bestimmt werden kann, gestaltet sich dies bei Baumsegmenten schwieriger. In den meisten Fällen ist eine exakte Konturlinie jedoch auch nebenrangig. Für ein Stadt-GIS z. B. genügt oft schon die Position des Stamms; evtl. ist noch eine Abschätzung der Kronenfläche von Interesse. Unter diesem Aspekt sollte auch eine Beurteilung der automatisiert ermittelten Segmentierungsergebnisse erfolgen. Die Erfassung einer Baumregion ist von größerer Bedeutung als die Übereinstimmung der Form mit der Vorgabe.

Die manuell ermittelte Vorgabe besteht aus 199 Baumsegmenten. Vereinzelt werden im automatisierten Ansatz reell vorhandene Bäume detektiert, die nicht in der Vorgabe enthalten sind (s. Abbildung 4.17, Markierung 1). Einerseits zeigt dies deutlich den Nutzen einer zumindest zu Kontrollzwecken hinzugezogenen automatisierten Segmentierung. Andererseits ist bei der Auswertung zu berücksichtigen, dass die nur automatisiert korrekt ermittelten Baumsegmente fälschlicherweise die Fehlalarmrate (s. nachfolgende Definition in Gleichung 4.7) erhöhen. Aus diesem Grund wird eine zusätzliche Messung unter Ausschluss dieser Regionen durchgeführt.

Bei der Gegenüberstellung von Vorgabe und Segmentierungsergebnis sind zwei besondere Situationen hervorzuheben. Im Schatten liegende Baumsegmente werden durch das vorgeschlagene Vorgehen meist nicht erkannt (s. Abbildung 4.17, Markierung 2). Zwar weisen sowohl der NDVI- als auch der Farbtonansatz durch die darin implementierten Normierungen eine gewisse Unabhängigkeit gegenüber Helligkeitsänderungen auf. Doch im Schatten sind die charakteristischen Spektralmerkmale durch den niedrigen Signalpegel nicht mehr klar ausgebildet. Beleuchtungseffekte und Rauschen haben relativ zum Signal größere Auswirkungen. Ähnlich verhält es sich mit nicht ausreichend beleuchteten oder weniger vitalen Bäumen. Der spektrale Unterschied zu ausreichend beleuchteter vitaler Vegetation ist in solchen Fällen zu groß, um nach den Kriterien der verwendeten Ansätze noch als Vegetation zu gelten.

Eine andere Problematik liegt bereits im Oberflächenmodell begründet. Einige Bäume, besonders solche, die direkt an ein Haus angrenzen, sind im DOM nicht oder nicht ausreichend erfasst. Bei der DOM-Generierung wird dort aus Mangel an einander zugeordneten Stereopunkten interpoliert (s. Diskussion in Abschnitt 2.5.2). Dies kann bei Extraktion erhöhter Regionen (s. Kapitel 3) zu zu kleinen Regionen führen. Die Verschneidung von zu kleinen erhöhten Regionen mit der Vegetationsseg-





**Abbildung 4.17:** Gegenüberstellung von Baumsegmentierung (grün) und manueller Vorgabe (rot). Anmerkungen zu den mit einer Ziffer versehenen Sonderfällen sind dem Text zu entnehmen.

mentierung resultiert dann ebenfalls in zu kleinen Segmenten. Sie weisen nicht die Mindestabmessungen eines Baumes auf und werden deshalb entfernt. Ein Bildbeispiel dafür ist in Abbildung 4.17 an Markierung 3 gezeigt.

In Tabelle 4.2 ist die Auswertung der Baumdetektion anhand der manuell ermittelten Vorgabe in Form von Erkennungsraten zusammengefasst. Bezeichnet man die Menge der manuell vorgegebenen Pixel oder Regionen als „SOLL“ und die der automatisiert ermittelten als „IST“, so sind die Detektionsrate (DR), deren Komplement, die Fehlrückweisungsrate (FRR), sowie die Fehlalarmrate (FAR) folgendermaßen definiert:

$$DR = \frac{|\text{IST} \cap \text{SOLL}|}{|\text{SOLL}|}; \quad FRR = \frac{|\overline{\text{IST}} \cap \text{SOLL}|}{|\text{SOLL}|}; \quad FAR = \frac{|\text{IST} \cap \overline{\text{SOLL}}|}{|\text{IST}|} \quad (4.7)$$

Die Erkennungsraten in Tabelle 4.2 werden einmal auf Basis der abgedeckten Fläche, das andere Mal auf Basis der Anzahl abgedeckter Regionen ermittelt. Bei der Zählung der abgedeckten Regionen spielt der Grad der Abdeckung der einzelnen (Baum-)Regionen keine Rolle.

Für eine vergleichende Gegenüberstellung der Erkennungsraten erfolgt die Baumsegmentierung jeweils separat mit dem NDVI- und Farbtonkriterium sowie kombiniert entsprechend dem Vorgehen

Grundlage	Maß	Detektions- rate (%)	Fehlrückweisungs- rate (%)	Fehlalarmrate (%)
NDVI	Fläche	57,09	42,91	21,17
	Regionen	82,41	17,59	43,93
Farbton	Fläche	64,21	35,79	20,13
	Regionen	85,93	14,07	42,72
NDVI und Farbton	Fläche	60,33	39,67	11,49
	Regionen	74,87	25,13	13,04
➡ mit veränderter Vorgabe	Fläche	82,50	17,50	10,53
	Regionen	83,59	16,41	6,47

**Tabelle 4.2:** Genauigkeitsangaben für die Baumdetektion (jeweils in Prozent).

in Abschnitt 4.2.3. Die Segmentierungen mit jeweils nur einem Kriterium erreichen bis auf einen Fall leicht höhere Detektionsraten als das kombinierte Vorgehen, jedoch zu Lasten der Fehlalarmrate. Es werden also deutlich mehr Segmente detektiert als in der Vorgabe vorhanden. Besonders deutlich wird dies bei Betrachtung der abgedeckten Regionen. Bei NDVI- bzw. Farbtonsegmentierung werden 164 bzw. 171 von 199 Segmenten erkannt, bei der kombinierten Vorgehensweise 149. Der Anteil falsch positiv detektierter Segmente beträgt bei NDVI- bzw. Farbtonsegmentierung 145/330 bzw. 132/309, bei der Kombination lediglich 24/184. Die Kombination der beiden Merkmale führt also hauptsächlich zu einer sinnvollen Bereinigung der detektierten Segmente.

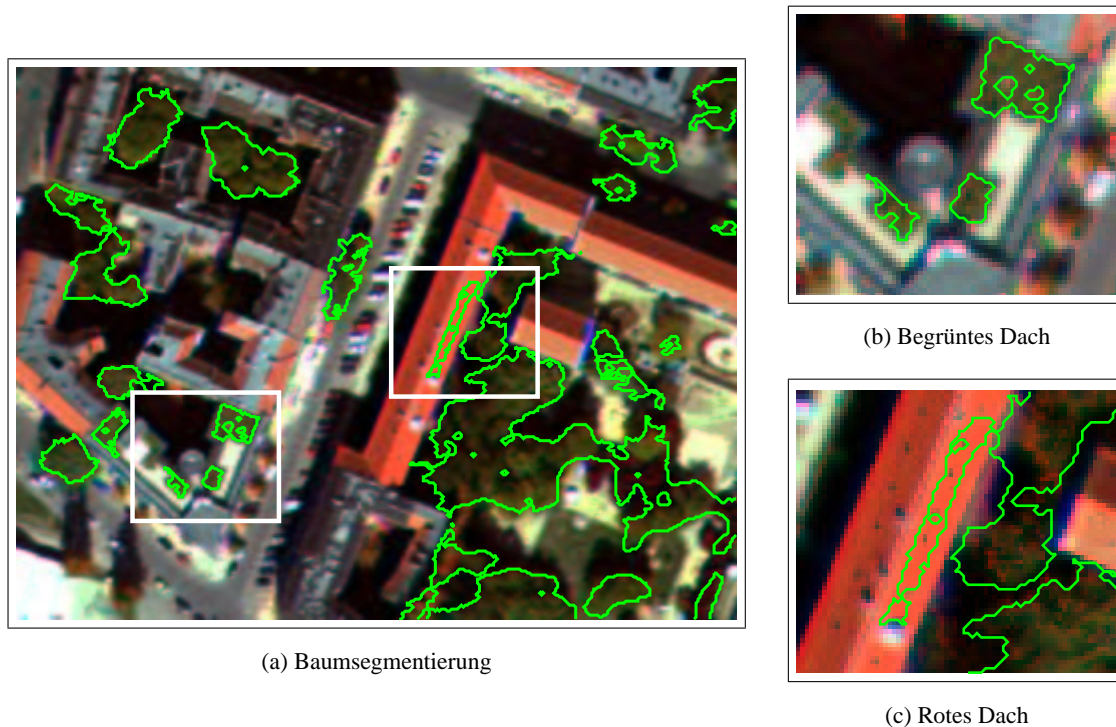
Wie bereits erwähnt, werden Bäume im Schatten nur schwer erkannt. Auch Bäume, denen im DOM fälschlicherweise eine Grundhöhe zugeordnet ist, können nicht detektiert werden. Aus diesem Grund wird die Vorgabe für eine weitere Messung um diese „nicht-erkennbaren Segmente“ reduziert. Weiterhin werden die bei einem regionenbasierten Vorgehen detektierten Segmente bereinigt von Regionen, die fälschlicherweise nicht in der Handsegmentierung enthalten sind. Resultat sind die in den beiden letzten Zeilen von Tabelle 4.2 wiedergegebenen Erkennungsraten. Die Detektionsrate liegt nun bei mehr als 80 %.

#### 4.4.3 Gebäudedetektion

Zur Gebäudedetektion wird entsprechend dem in diesem Kapitel herausgearbeiteten Vorgehen verfahren. Die Vegetationssegmente werden anhand des roten Farbtons und des NDVI-Wertes bestimmt. Die resultierenden Segmente werden morphologisch nachverarbeitet. Von den verbleibenden erhöhten Regionen werden Schattenpixel an der Segmentgrenze entfernt. Aus einer morphologischen Nachverarbeitung dieser Regionen ergeben sich dann die endgültigen Gebäudesegmente.

Bei der Beurteilung der Qualität der Trennung von Vegetation und Gebäuden sind neben klaren Fehlzuordnungen zwei Sonderfälle herauszustellen. Zum einen begrünte Dächer oder Dachgärten,





**Abbildung 4.18:** Fehlzusammenordnungen in der Baumsegmentierung. Farbkomposit mit überlagerten Konturen aus der Segmentierung erhöhter Vegetation (a) und Ausschnittsrahmen für die Beispiele (b) und (c).

die aufgrund der spektralen Charakteristika der Vegetationsklasse zugeordnet werden (s. Bildbeispiel in Abbildung 4.18(b)). Da optische Fernerkundungssensoren nur die Oberfläche der Objekte erfassen, weisen begrünte Dächer die gleichen spektralen Merkmale wie Vegetation auf. Eine Unterscheidung alleine durch die spektrale Information ist in diesem Fall also nicht möglich. Jedoch könnten im Fall von begrünten Dächern geometrische Merkmale als Unterscheidungskriterium dienen. Die reguläre, meist rechtwinklige Dachform unterscheidet sich von der „ausgefranst“ Konturlinie der Baumsegmente.

Einen anderen Sonderfall stellen Segmente dar, die in der in Abschnitt 4.2.2 gewählten Kanalkombination einen ähnlichen Farbtönen aufweisen wie Vegetation. Dies ist hauptsächlich bei ausgewählten Ziegeldächern, die in einem bestimmten Winkel beleuchtet sind, der Fall. Aus diesem Grund werden in Abschnitt 4.2.3 von den roten Segmenten nur diejenigen als Vegetation eingestuft, die auch einen genügend hohen NDVI-Wert aufweisen. Problematisch ist bei diesem Vorgehen der Fall eines Baumes, der direkt an ein rotes Hausdach angrenzt. Dach und Baum bilden dann *ein* Segment. Ist der Vegetationsanteil darin so hoch, dass der NDVI-Wert für das gemeinsame Segment über dem gewählten NDVI-Schwellwert liegt, gilt das Segment als Vegetation. Somit fehlt der entsprechende Dachanteil in der Gebäudesegmentierung. In Abbildung 4.18(c) ist ein Bildbeispiel zu dieser Proble-

matik dargestellt.

Die in der manuellen Segmentierung erfassten Wohnblocks sind Altbauten, die ein Flachdach mit Dachschräge zur Außenkontur hin haben. Diese Dachschrägen können zu einem Problem für die sehr einfach implementierte Schattendetektion werden. Während ebene Dachflächen meist keinem Schattenwurf ausgesetzt sind (s. Diskussion in Abschnitt 4.3.1), ist dies bei Dachschrägen durchaus der Fall. Da sich beschattete Dachschrägen zudem am Rand des Gebäudesegments befinden, werden sie bei der Elimination der Schattenpixel entfernt. Dies ist nur durch eine ausgefeiltere Schattendetektion zu beheben. Da die Dachschrägen jedoch nur einen kleinen Teil der Gebäudefläche ausmachen, fällt dieses Problem für die Gebäudedetektion nicht zu sehr ins Gewicht<sup>16</sup>.

In der automatisierten Vorgehensweise wird gegenüber der manuellen Segmentierung ein kleineres Gebäude nicht erfasst. Weiterhin werden bei acht von 89 Hinterhöfen die entsprechenden Segmentgrenzen nicht erkannt. Da in diesem Fall für die Gebäudesegmentierung der Hinterhof nicht existent ist, erstreckt sich das Gebäude auch über den Hof. Die Ursache dafür liegt im Höhenmodell, denn an den betreffenden Stellen reicht das DOM nicht auf den Grund. Wie bei dem analogen Problem in der Baumdetecktion wird bei der DOM-Generierung an diesen Stellen aus Mangel an einander zugeordneten Stereopunkten interpoliert (s. Diskussion in Abschnitt 2.5.2).

Ein Vergleich der automatisiert erkannten Gebäudepixel mit der manuellen Vorgabe führt zu den in Tabelle 4.3 dargestellten Detektionsraten. 91,5 % der automatisiert erkannten Gebäudepixel stimmen mit der Vorgabe überein, 8,5 % der Sollvorgabe hingegen werden nicht als Gebäude eingestuft. Letzteres spiegelt sich in lediglich einem nicht erkannten Gebäude sowie Abweichungen in der Form der Gebäudesegmente wider. Die Fehlalarmrate, also der Anteil der automatisiert ermittelten Gebäudepixel, die nicht der manuellen Vorgabe entsprechen, liegt bei 10 %. Dies macht sich nicht durch in Realität nicht vorhandene Häuser bemerkbar, sondern lediglich durch Abweichungen in der Form der Gebäudesegmente.

Gebäuderegion	DR (%)	FRR (%)	FAR (%)
Manuelle Vorgabe	91,51	8,49	9,95
➡ nur erkannte Hinterhöfe	91,51	8,49	8,68

**Tabelle 4.3:** Genauigkeitsangaben für die Gebäudedetektion.

Diese Auswertung schließt auch die nicht als Hinterhöfe erkannten Regionen ein. Da die Ursache für dieses Nicht-Erkennen jedoch bereits im DOM liegt, erfolgt auch eine Auswertung unter Ausschluss der betroffenen Hinterhöfe. Dazu werden in der automatisiert gewonnenen Gebäudesegmentierung die entsprechenden Hinterhöfe entfernt. Da die manuelle Vorgabe nicht verändert

<sup>16</sup>Bei der in Kapitel 6 behandelten Optimierung der Gebäudekontur wirkt sich das Fehlen von Randsegmenten wie Dachschrägen allerdings aus. Die Adaption der Gebäudekontur kann sich dann an den falschen Bildkanten orientieren, z. B. dem Dachfirst.

wird, bleiben die Detektions- und Fehlrückweisungsrate gleich. Die Fehlalarmrate hingegen sinkt auf 8,68 %.



## 5 Segmentierung in einzelne Objekte

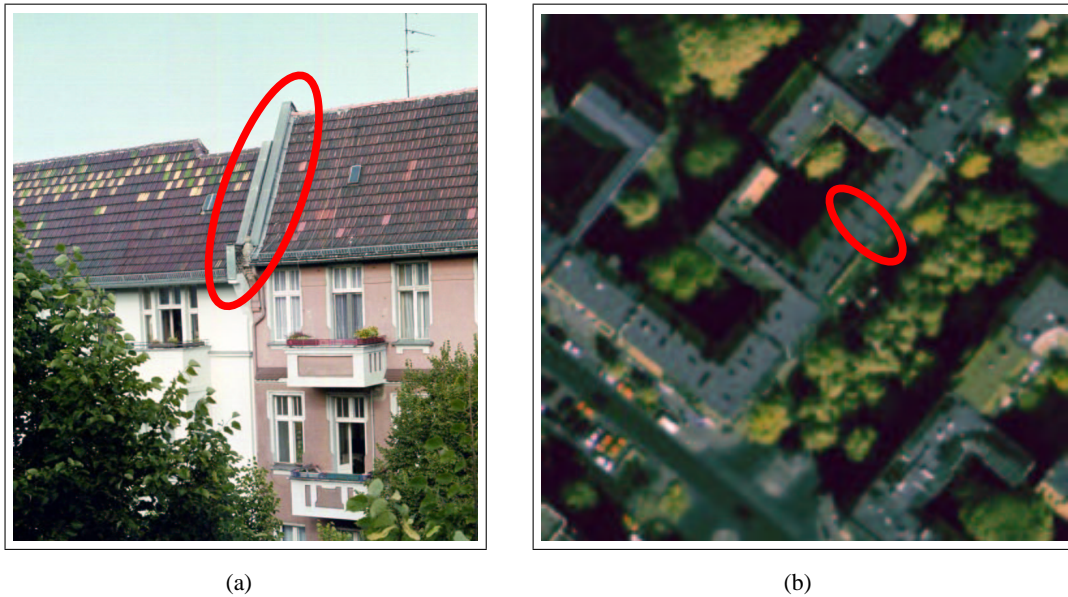
Im vorangegangenen Kapitel wurden die erhöhten Regionen in Gebäude- und Vegetationspixel segmentiert. Ein solches Segment kann aus einem Objekt wie auch aus einer Vielzahl zusammenhängender Objekte bestehen. So kann ein Vegetationssegment aus einem einzigen Baum wie auch aus einem Waldstück bestehen, ein Gebäudesegment kann ein einzelnes Haus wie auch ein gesamter Häuserblock sein. Ziel dieses Kapitels ist es, die Segmente in einzelne Objekte aufzuteilen. Damit wird zum einen die Objektkomplexität reduziert, zum anderen wird damit die in Abschnitt 1.4.2 definierte Objektebene, die unterste Ebene im vorgeschlagenen hierarchischen Vorgehen zur Objektextraktion, erreicht. Die Erfassung von Gebäudedaten in einem Stadt-GIS erfolgt in der Regel mindestens auf Ebene einzelner Häuser, durchaus auch detailreicher. Das liegt daran, dass man erst ab dieser Stufe Eigentumsverhältnisse zuweisen und somit auch rechtliche Einheiten erfassen kann.

In diesem Kapitel werden ausschließlich Gebäudesegmente unterteilt in kleinere Einheiten, möglichst einzelne Häuser. Zur Unterteilung von Baumgruppensegmenten in einzelne Bäume sei verwiesen auf die Ansätze von PINZ (1989), GOUGEON (1995), BRANDTBERG UND WALTER (1998), WULDER ET AL. (2000), POULIOT ET AL. (2002) sowie auf den von POLLOCK (1996) eingeführten und von LARSEN UND RUDEMO (1997) fortentwickelten modellbasierten Ansatz.

Die Unterteilung der Gebäudesegmente orientiert sich hauptsächlich an der Gradienteninformation eines Orthobildes. Starke Grauwertwechsel deuten auf eine Kante innerhalb des Segments hin. Jedoch führen meist — insbesondere bei Gebäudesegmenten — auch zahlreiche Teile eines Objekts (z. B. Dachantennen) zu hohen Grauwertänderungen. Zur Differenzierung echter Unterteilungen von lokalen Grauwertsprüngen müssen deshalb weitere Merkmale wie die Form der Unterteilung oder die Grauwertverteilung entlang der Unterteilung hinzugezogen werden.

Nach einer Einführung in die Thematik und einem Literaturüberblick in Abschnitt 5.1 werden in Abschnitt 5.2 die Vorverarbeitungsschritte des Ansatzes zur Unterteilung von Gebäudesegmenten vorgestellt.

In Abschnitt 5.3 wird die Auswahl geeigneter Ausgangspunkte für eine linienhafte Unterteilung der Gebäudesegmente behandelt. Die resultierenden Ausgangspunkte führen mit der ihnen zugeordneten Orientierung zu Kandidaten auf der Basis von Trennlinien. In den Abschnitten 5.4 bis 5.6 werden diese Linien aufgrund geometrischer Kriterien selektiert und anhand eines Gütekriteriums bewertet. Eine konsistente Unterteilung der Gebäudesegmente durch die Trennlinien erfolgt dann in Abschnitt 5.7 anhand des Gütekriteriums. Resultat sind Regionen, die idealerweise einzelnen Häusern entsprechen.



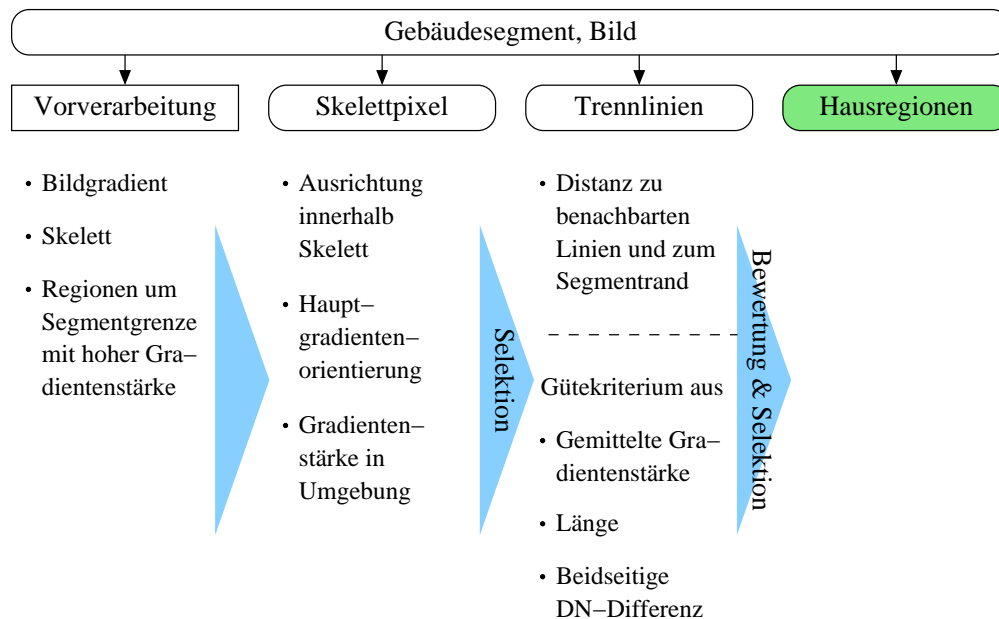
**Abbildung 5.1:** Brandmauern am Übergang zweier aneinandergrenzender Häuser; einmal aus der Seitenansicht (a), das andere Mal aus der Orthoperspektive (b). Die rot überlagerten Ellipsen geben die korrespondierende Brandmauer in beiden Bildern wieder.

Dieses Kapitel endet mit einer Auswertung der vorgestellten Algorithmen in Bezug auf die damit erzielbare Genauigkeit und die benötigte Rechenzeit (Abschnitt 5.8).

### 5.1 Einführung und Literaturüberblick

Aus der Orthoperspektive lassen sich die Dächer einzelner Häuser eines Gebäudekomplexes gut unterscheiden, wenn sie aus verschiedenen Dachmaterialien bestehen. So sind benachbarte Dächer, die jeweils mit roten und schwarzen Ziegeln gedeckt sind, visuell sehr gut voneinander zu unterscheiden. Auch eine unterschiedliche Textur der Dachmaterialien, andere Gebäudehöhen oder verschiedene Dachneigungen tragen zur Differenzierbarkeit aneinandergrenzender Häuser bei. Doch von einer Andersartigkeit dieser Merkmale kann nicht immer ausgegangen werden. Als einziges, auch baulich zwingend vorhandenes Merkmal zur Unterscheidung benachbarter Häuser aus der Orthoperspektive verbleiben die *Brandmauern*. Ab einer geometrischen Auflösung von etwa einem halben Meter sind diese Trennungswände auch deutlich in einer Luftaufnahme sichtbar. Abbildung 5.1 zeigt dieselbe Brandmauer einmal in der Seitenansicht sowie aus der Orthoperspektive. Selbst bei sonst identischen Dachstrukturen ist diese Trennungslinie auszumachen.

In diesem Kapitel wird ein Vorgehen zur Unterteilung von Gebäudesegmenten vorgestellt, das sich hauptsächlich an den durch Brandmauern hervorgerufenen linienhaften Bildkanten (hier bezeichnet als „Trennlinie“) orientiert. Zu diesem Vorgehen gibt es nur bedingt vergleichbare Ansätze in der



**Abbildung 5.2:** Schematische Darstellung zur Unterteilung der Gebäudesegmente in kleinere Einheiten.

veröffentlichten Literatur. Ein verwandtes Gebiet ist die Detektion von Linien im Hinblick auf eine spätere Gebäudeextraktion. Der Liniendetektor von BURNS ET AL. (1986) wird von zahlreichen Autoren für diese Zwecke eingesetzt (s. z. B. YOON ET AL. 1999). Dessen Vorgehen beruht auf der Bestimmung von Linienregionen aus der Gradientenorientierung. Eine solche Region besteht aus Pixeln mit jeweils entgegengesetzter Orientierung (beide Seiten der Linie). Zu jeder geradlinigen Bildkante wird so eine optimale Region zugehöriger Pixel bestimmt. In diese Region wird schließlich eine Linie eingepasst.

Als anderer Ansatz zur Detektion gerader Bildkanten sei hier nur die Hough-Transformation genannt (s. O’GORMAN UND CLOWES 1976). Das in diesem Kapitel vorliegende Problem, Gebäudesegmente durch Trennlinien zu unterteilen, unterscheidet sich jedoch in mehreren Punkten von einer allgemeinen Liniendetektion: Aus *Startpunkten* (geeignete Skelettpixel) und einer Abschätzung der *Linienorientierung* ergibt sich durch die Segmentgrenzen auch gleichzeitig die Linienlänge. Auf diese Weise können schnell *Hypothesen für Trennlinien* aufgestellt und anhand von weiteren Kriterien überprüft werden. Ferner gilt es, zusammenhängende Unterteilungslinien zu extrahieren. Auch wenn die Gradientenstärke über die gesamte Linie gesehen nicht gleichmäßig groß ist, soll *eine einzige Linie* und nicht mehrere Liniensegmente resultieren.

Abbildung 5.2 gibt einen Überblick über den in drei Verarbeitungsebenen gegliederten Ablauf der Gebäudeunterteilung (s. auch MAYER 2002). Durch die Anwendung der im letzten Kapitel genannten Methoden stehen Segmente von Gebäuden zur Verfügung. Zusammen mit dem Orthobild eines Bildkanals, vorzugsweise des Nadirs, stellen diese die Eingangsdaten für die Unterteilung dar. In

einem Vorverarbeitungsschritt werden daraus der Bildgradient sowie die zentralen Skelettpixel der Gebäudesegmente berechnet. Am Rand der Gebäudesegmente werden dann Regionen hoher Gradientenstärke bestimmt. Anhand der Gradientenorientierung der Pixel dieser Regionen wird letztlich die Ausrichtung der Trennlinien bestimmt. Nach diesem Vorverarbeitungsschritt werden die Gebäudesegmente in den nun folgenden Schritten segmentweise abgearbeitet. Die ursprünglich sehr hohe Zahl an Kandidaten für eine Unterteilung wird mit jedem Arbeitsschritt aufgrund des aktuell betrachteten Kriteriums reduziert. Die ersten Tests sind einfach und dienen einer schnellen Vorauswahl. Die späteren Überprüfungen werden aufwändiger, aber dafür sind bereits viele der Kandidaten ausgeschlossen worden. Die schrittweise Selektion der Kandidaten erfolgt in drei Bearbeitungsebenen. Zuerst auf Grundlage der *Skelettpixel*, dann basierend auf den daraus abgeleiteten *Trennlinien* und schließlich anhand der so entstandenen *Hausregionen*.

### 5.2 Begriffsdefinition und Vorverarbeitungsschritte

Da die mit Brandmauern verbundenen „Trennlinien“ im Zentrum des präsentierten Vorgehens zur Gebäudeunterteilung stehen, wird dieser Begriff zunächst näher definiert:

**Definition 5.1** *Eine Trennlinie unterteilt ein Gebäudesegment in zwei Teilsegmente, weist eine hohe mittlere Gradientenstärke bei möglichst konstanter Gradientenorientierung auf und verläuft möglichst senkrecht zu den beiden von der Trennung betroffenen Seiten des Gebäudesegments.*

Diese Definition spiegelt nicht vollkommen die Realität wider. Brandmauern verlaufen zwar meist senkrecht zur Gebäudefassade, doch ist das keineswegs zwingend. Außerdem ist es bei komplexeren Anordnungen von Häusern innerhalb eines Häuserblocks möglich, dass Brandmauern auch innerhalb des Gebäudesegments enden<sup>1</sup>. Diese relativ geringfügigen Einschränkungen im Vergleich zu real vorkommenden Brandmauern stellen den Rahmen für das vorgestellte Vorgehen dar. Um einen visuellen Eindruck von dem trennlinienbasierten Vorgehen zur Gebäudeunterteilung zu erhalten, sei bereits jetzt auf die Darstellung von Trennlinien in den Abbildungen 5.13–5.16 verwiesen.

Die Generierung von Trennlinien basiert zu einem wesentlichen Teil auf dem *Gradienten* des zu bearbeitenden Bildes. Er wird nach Gleichung 1.1 in Abschnitt 1.5.3 aus einem Bildkanal, z. B. dem Nadir, berechnet. Mit dem Gradienten steht einerseits die Gradientenstärke  $|\nabla|$  und andererseits die Gradientenorientierung  $\varphi$ , also die maßgebliche Kantenrichtung für jedes Pixel, zur Verfügung.

Die Ausrichtung einer Trennlinie wird durch die Gradientenorientierung der Pixel mit vorab festgelegtem Abstand zum Segmentrand bestimmt. Allerdings macht es nur Sinn, die Orientierung von Pixeln mit hinreichend hoher Gradientenstärke zu berücksichtigen. Aus diesem Grund werden vorab entlang der Gebäudesegmentgrenze *Regionen hoher Gradientenstärke*, hier bezeichnet mit  $\mathcal{G}$ , bestimmt.

---

<sup>1</sup> Auf diesen Aspekt wird in Abschnitt 5.7 nochmals genauer eingegangen.



Da die Unterteilung eines Gebäudesegments gemäß Definition 5.1 von einer Segmentseite zur gegenüberliegenden erfolgen soll, kreuzt die Trennlinie immer die Mittelachse des Segments. Aus diesem Grund basiert die initiale Definition von Trennlinien in dem vorgeschlagenen Vorgehen auf der Mittelachse. Geeignete Trennlinien sind über einen Punkt der Mittelachse sowie eine noch festzulegende Orientierung bestimmt. Die Berechnung der Segmentmittelachse wird in der Bildverarbeitung als Skelettieren bezeichnet. Dazu wird für jedes Segmentpixel der Abstand zur Segmentgrenze berechnet (TORIWAKI UND FUKUMURA 1978, BALLARD UND BROWN 1982, PAVLIDIS 1982, XIA 1989). In dem resultierenden „Entfernungsgebirge“ gilt die Mittelachse als der Grat. Nach dem Verfolgen dieses Grats maximaler Distanz werden die resultierenden *Skeletpixel* z. B. bei TORIWAKI UND FUKUMURA (1978) auf die Linienbreite von einem Pixel ausgedünnt.

Diese drei Vorverarbeitungsprodukte, der Bildgradient, die Regionen um die Segmentgrenze und die Skeletpixel, sowie als Ausgangsdaten ein Orthobild mit zugehöriger Gebäudesementierung sind die Basis für die nachfolgend beschriebenen Verarbeitungsschritte.

### 5.3 Trennlinien durch ausgewählte Skeletpixel

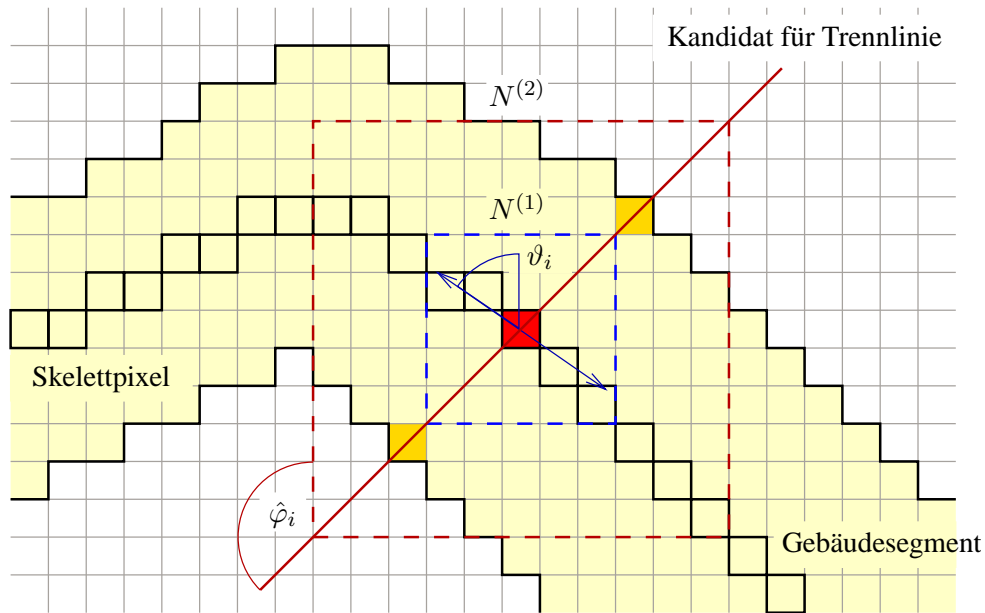
Wie bereits erwähnt, dienen die Skeletpixel als Ausgangspunkte für Trennlinien. In diesem Abschnitt wird ein Verfahren beschrieben, die Anzahl der Ausgangspunkte sinnvoll zu reduzieren und den Punkten eine Orientierung zuzuweisen — somit also Kandidaten für Trennlinien zu definieren. Zur Veranschaulichung der einzelnen Schritte dient die Darstellung in Abbildung 5.3.

Zuerst wird innerhalb einer kleinen Nachbarschaft  $N^{(1)}$  um jedes Skeletpixel mittels linearer Regression dessen *Ausrichtung*  $\vartheta$  *innerhalb des Skeletts* ermittelt. Der Winkel  $\vartheta$  wird relativ zur Bildsenkrechten bestimmt. Sollte sich die Skelettlinie in mehrere Linien auffächern, werden die Kreuzungspixel des Skeletts bei der Regression nicht miteinbezogen.

In einer größeren Nachbarschaft  $N^{(2)}$  um das aktuell betrachtete Skeletpixel  $i$  wird die *Hauptgradientenorientierung*  $\hat{\varphi}$  *von Segmentrandpixeln* mit ausreichend hoher Gradientenstärke ermittelt.  $N^{(2)}$  muss groß genug sein, um die Segmentgrenzen zu erreichen. Die Regionen hoher Gradientenstärke entlang der Segmentkontur stammen aus der Vorverarbeitung. Für das aktuelle Pixel seien diese Randpixel in der Menge  $\mathcal{G}_i$  erfasst. Die dominante Gradientenorientierung ergibt sich durch Betrachtung der Häufigkeiten  $\mathbf{h}$  in einem durch Mittelwertfilterung geglätteten Histogramm der Gradientenorientierungen (analog zu dem Vorgehen in Abschnitt 3.3):

$$\hat{\varphi}_i = \operatorname{argmax}_{\varphi_j} \mathbf{h}(\varphi_j \bmod \pi), \quad \{j | j \in \mathcal{G}_i\}. \quad (5.1)$$

Die Gradienten gegenüberliegender Gebäudeänder sind im Normalfall gegenläufig orientiert. Aus diesem Grund wird in Gleichung 5.1 die Hauptorientierung durch die Modulo-Operation auf zwei Quadranten reduziert. So tragen auch um  $\pi$  versetzte Gradientenorientierungen zum gleichen Histogrammeintrag bei. Abbildung 5.4 veranschaulicht die histogrammbasierte Bestimmung von  $\hat{\varphi}$  an ei-



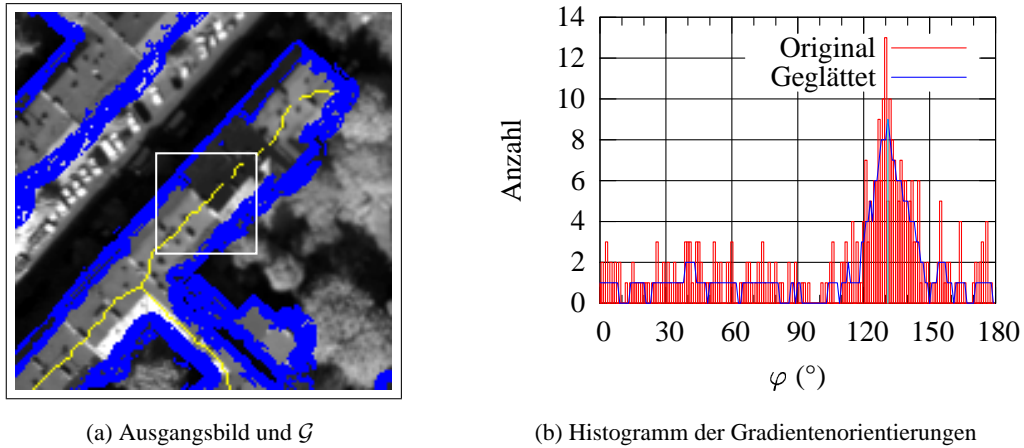
**Abbildung 5.3:** Schematische Darstellung zur Entwicklung eines Trennlinienkandidaten. Der Ausschnitt aus einem Gebäudesegment und die zugehörigen Skelettpixel sind in hellem Gelb dargestellt. Um das aktuell betrachtete Skelettpixel (rot) wird innerhalb einer Nachbarschaft (blau gestrichelt) die Ausrichtung der Skelettpixel (dunkelblauer Doppelpfeil) sowie die mittlere Gradientenstärke ermittelt. Die Hauptgradientenorientierung wird anhand der Segmentrandpixel im rot gestrichelten Fenster berechnet. Die Endpunkte der Trennlinie (orange) ergeben sich aus einem Schnitt mit den Segmentgrenzen.

nem Bildbeispiel. Aus den blau dargestellten Segmentrandpixeln hoher Gradientenstärke des inneren Ausschnitts von Abbildung 5.4(a) ergibt sich das Histogramm in Abbildung 5.4(b). Das Maximum und somit  $\hat{\varphi}$  ist in diesem Beispiel bei  $131^\circ$ .

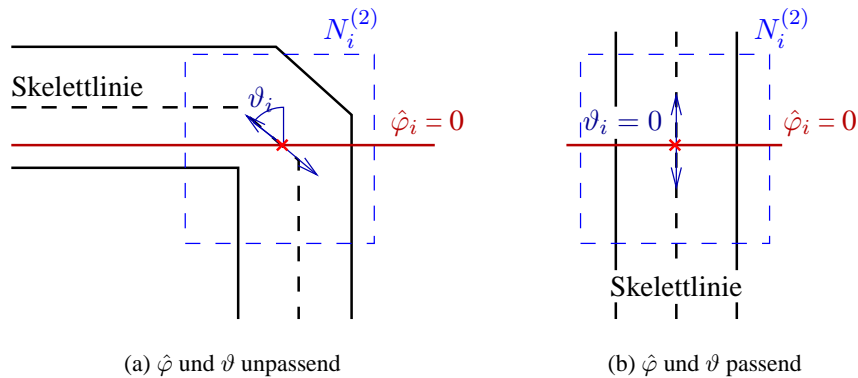
Anhand der Gradientenorientierung  $\hat{\varphi}$  wird die Ausrichtung der Trennlinie bestimmt. Die Trennung in zwei Teilsegmente erfolgt senkrecht zum Gebäuderand, also in Richtung  $\hat{\varphi}$ . Falls das Gebäudesegment, hier repräsentiert durch die Skelettpixel, jedoch eine dazu unpassende Ausrichtung aufweisen sollte, wird das aktuelle Skelettpixel als Basispunkt für eine Trennlinie verworfen. Diese Entscheidung erfolgt aufgrund der Winkeldifferenz  $|\hat{\varphi}_i - \vartheta_i|$ . Abbildung 5.5 veranschaulicht diesen Selektionsschritt an zwei Fallbeispielen.

Den verbleibenden Skelettpixeln wird die jeweilige Gradientenorientierung  $\hat{\varphi}$  als Trennlinienorientierung zugeordnet. Da die Häuser eines Gebäudesegments oftmals gleich ausgerichtet sind, werden viele Skelettpixel eine ähnliche zugeordnete Trennlinienorientierung aufweisen. Um die Trennlinienorientierungen innerhalb eines Segmentes zu vereinheitlichen, werden *Gruppen ähnlicher Orientierungen* ermittelt. Für jede dieser Gruppen werden die Orientierungen auf eine Hauptorientierung reduziert.

Eine letzte Selektion auf „Skelettpixelebene“ basiert auf der Gradientenstärke um das aktuell be-



**Abbildung 5.4:** Bestimmung von  $\hat{\varphi}$  durch Betrachtung des Histogramms der Gradientenorientierungen (b). Im zugehörigen Ausgangsbild (a) sind die Skelettpixel gelb und die Segmentrandpixel hoher Gradientenstärke blau dargestellt.



**Abbildung 5.5:** Schematische Darstellung des Zusammenhangs von lokaler Skelettpixelausrichtung  $\vartheta$  und Hauptgradientenorientierung  $\hat{\varphi}$ . Für das rot markierte aktuelle Pixel  $i$  wird  $\vartheta_i$  sowie durch die Randpixel in  $N_i^{(2)}$   $\hat{\varphi}_i$  bestimmt. Ist die Skelettlinie nicht annähernd senkrecht zur Ausrichtung der Trennlinie (gegeben durch  $\hat{\varphi}_i$ ), wird  $i$  als Kandidat für eine Trennlinie verworfen (Fall (a)). Andernfalls (b) wird ein Trennlinienkandidat generiert.

trachtete Pixel  $i$ . Da die Trennlinien an Bildkanten realisiert werden sollen, muss auch für jeden Punkt der Linie — somit auch für  $i$  — eine hinreichend hohe Gradientenstärke vorhanden sein. Deshalb werden Skelettpixel mit geringem  $|\nabla|$  als Basispunkte für eine Trennlinie ausgeschlossen. Um mögliche Lokalisierungsfehler zu kompensieren wird die Gradientenstärke in einer kleinen Nachbarschaft  $N_i^{(1)}$  um  $i$  gemittelt. Da  $i$  mittlerweile eine Trennlinienorientierung zugeschrieben wurde, können bei der Mittelung Pixel mit unpassender Gradientenorientierung ausgeschlossen bzw. durch einen Gewichtungsfaktor unterdrückt werden. Über die absolute Differenz von Hauptorientierung zu der Orientierung eines Pixels  $j$  in der Umgebung  $N_i^{(1)}$ ,  $\Delta\varphi_{ij}$ , ergibt sich die *gewichtet gemittelte Gradientenstärke*  $|\overline{\nabla}|$ :

$$\Delta\varphi_{ij} = \min(|\hat{\varphi}_i - (\varphi_j \bmod \pi)|, \pi - |\hat{\varphi}_i - (\varphi_j \bmod \pi)|), \quad (5.2)$$

$$|\overline{\nabla}|^{N_i^{(1)}} = \frac{1}{|N_i^{(1)}|} \sum_{j \in N_i^{(1)}} |\nabla|_j \left(1 - \frac{\min(\Delta\varphi_{ij}, \Delta\varphi_{\max})}{\Delta\varphi_{\max}}\right). \quad (5.3)$$

Dabei ist eine maximal zulässige Winkeldifferenz  $\Delta\varphi_{\max}$  zu spezifizieren. Ist  $\Delta\varphi_{ij}$  kleiner als  $\Delta\varphi_{\max}$  fließt die Gradientenstärke des Pixels  $j$  in die Mittelung ein. Die Gewichtung der Gradientenstärke mit einem Faktor im Intervall  $[0, 1]$  erfolgt linear in Abhängigkeit von  $\Delta\varphi_{ij}$ .

Durch die Orientierung zu jedem der verbleibenden Basispunkte (in Form von Skelettpixeln) sind die Trennlinien definiert. Reduziert man diese Geraden auf die beidseitig jeweils nächsten Segmentgrenzen, resultieren Liniensegmente bzw. Strecken<sup>2</sup> der *Länge*  $l_i$  daraus. Zu kurze Linien, also mit  $l_i < l_{\min}$ , werden von der weiteren Bearbeitung ausgeschlossen.

## 5.4 Lage der Linien im Gebäudesegment

Bevor die Trennlinienkandidaten im nächsten Abschnitt bewertet werden, erfolgt eine Vorauswahl und Modifikation basierend auf deren Lage im Gebäudesegment. Ziel ist das Unterdrücken ungeeigneter Linien, die jedoch eine gute Bewertung erzielen können.

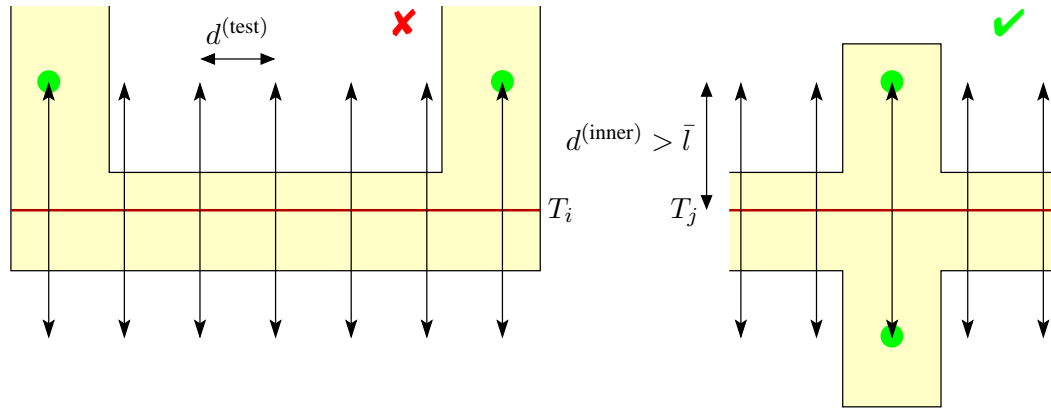
### 5.4.1 Ausschluss von Linien ohne beidseitige Segmentfortsetzung

In einem ersten Schritt wird überprüft, ob sich das Gebäudesegment sowohl auf der einen wie auch auf der anderen Seite der Trennlinie an mindestens einer Stelle fortsetzt. Denn nur dann kann davon ausgegangen werden, dass die Linie nicht nur einen marginalen Teil des Segments „abschneidet“, sondern zwei größere Einheiten unterteilt.

Wie in Abbildung 5.6 dargestellt, werden dazu im Abstand  $d^{(\text{test})} = l_{\min}$  beginnend in der Mitte der Trennlinie Stichpunkte überprüft. Diese Stichpunkte sind jeweils  $d^{(\text{inner})}$  von der Trennlinie

---

<sup>2</sup>Im Folgenden wird weiterhin von Trennlinien gesprochen. Gemeint sind jedoch immer die soeben definierten Liniensegmente oder Strecken.



**Abbildung 5.6:** Schematische Darstellung zum Ausschluss von Linien, für die es keine beidseitige Fortsetzung im Gebäudesegment gibt. Die links dargestellte Trennlinie  $T_i$  wird ausgeschlossen, da kein Stichpunktpaar komplett im Gebäudesegment liegt. Die rechte Linie  $T_j$  hingegen gilt als weiterer Trennlinienkandidat.

entfernt (in Abb. 5.6 sind die Stichpunkte durch die Pfeilenden symbolisiert). Um zu testen, ob sich das Gebäudesegment an einem der Stichpunkte beidseitig weiter ausdehnt, muss sich *mindestens ein Stichpunktpaar komplett im Gebäudesegment* befinden. Ferner muss der Abstand  $d^{(\text{inner})}$  so gewählt werden, dass er größer als die mittlere Hausbreite  $\bar{l}$  ist. Linien, für die nicht wenigstens ein Stichpunktpaar im Gebäudesegment liegt, werden verworfen.

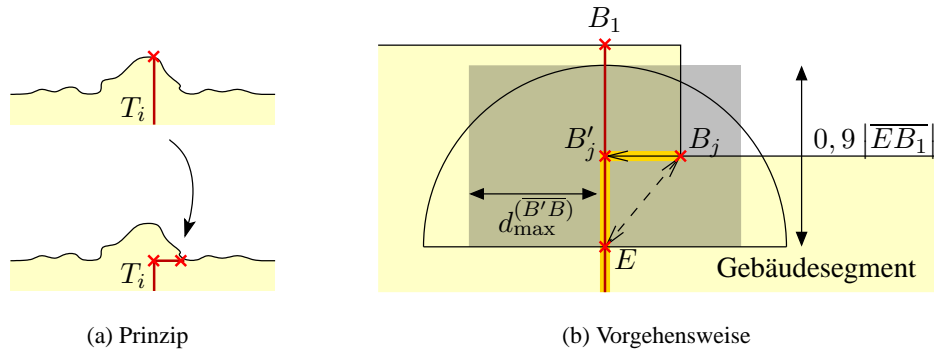
### 5.4.2 Verkürzung der Linienenden

Durch den im letzten Kapitel beschriebenen Segmentierungsprozess werden die Gebäudekonturen nicht immer einwandfrei erfasst. Die Segmentränder erscheinen in vielen Fällen „ausgefranst“. Trennlinien, die in einer „Ausbeulung“ der Segmentkontur liegen (s. Abbildung 5.7(a)), werden nun so modifiziert, dass an den Enden der Linie jeweils der kürzeste Weg zum Segmentrand gewählt wird. Dazu werden im Bedarfsfall zwei zusätzliche Stützpunkte eingeführt. Zur Verdeutlichung der im Folgenden beschriebenen Vorgehensweise dient Abbildung 5.7(b) (dabei wird nur die eine Seite zwischen Skelettpixel und Gebäuderand betrachtet).

Um die gebäudetypische Rechtwinkligkeit zu fördern, erfolgt eine Linienverkürzung ausschließlich orthogonal zur Richtung der Trennlinie. Die Strecke vom zentralen Skelettpixel  $E$  zum Segmentrand  $B_1$  sei  $\overline{EB_1}$ . Aufgrund der Orthogonalitätseinschränkung ist ein anderes Segmentrandpixel  $B_j$  von  $E$  aus nur über die orthogonale Projektion von  $B_j$  auf die Trennlinie,  $B'_j$ , zugänglich. Die Länge des Weges von  $E$  nach  $B_j$  über  $B'_j$  sei  $d^\perp(E, B_j)$  mit

$$d^\perp(E, B_j) = \left| \overline{EB'_j} \right| + \left| \overline{B'_j B_j} \right|. \quad (5.4)$$

Eine Verkürzung der Trennlinie liegt dann vor, falls  $d^\perp(E, B_j) < \left| \overline{EB_1} \right|$ . Um die Geradlinigkeit



**Abbildung 5.7:** Schematische Darstellung zur Verkürzung der Trennlinien an deren Enden. Links (a) der typische Einsatz an „ausgebeulten“ Segmenträndern. Das Vorgehen ist folgendermaßen (b): Innerhalb der dunkler unterlegten Region wird nach Segmentrandpixeln  $B_j$  gesucht. Ist die Distanz vom Skelettpixel  $E$  über  $B'_j$  kürzer als  $0,9 |\overline{EB_1}|$ , gilt  $B_j$  als Verkürzungskandidat.

einer geringfügigen Verkürzung vorzuziehen, wird dies modifiziert zu:

$$d^\perp(E, B_j) < 0,9 \cdot |\overline{EB_1}|. \quad (5.5)$$

Für die Strecke  $\overline{B'_j B_j}$  wird eine maximal erlaubte Länge  $d_{\max}^{(B'B)}$  festgesetzt, um sicherzustellen, dass sich die Trennlinie durch die Verkürzung nicht zu sehr verändert:

$$|\overline{B'_j B_j}| < d_{\max}^{(B'B)}. \quad (5.6)$$

Nur Werte  $d_{\max}^{(B'B)} < 0,9 |\overline{EB_1}|$  führen zu einer sinnvollen Begrenzung (in Abbildung 5.7 ist  $d_{\max}^{(B'B)} = 2/3 |\overline{EB_1}|$ ).

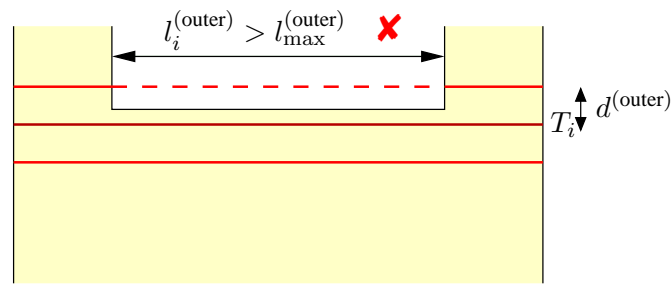
Aus diesen Randbedingungen ergibt sich folgendes Vorgehen. Ausgehend vom Basispunkt (Skelettpixel) der Trennlinie wird zu beiden Seiten der Linie eine Suchregion mit den Abmessungen  $2d_{\max}^{(B'B)} \times 0,9l_i$  aufgespannt<sup>3</sup>. Falls Segmentrandpixel  $B_j$  innerhalb dieser Region existieren, wird  $d^\perp(E, B_j)$  berechnet:

$$B_{\text{closer}} = \operatorname{argmin}_{B_j} d^\perp(E, B_j) \quad (5.7)$$

Ein neuer Polygonstützpunkt  $B'_{\text{closer}}$  wird eingeführt und das Linienende von  $B_1$  auf  $B_{\text{closer}}$  verschoben.

Eine Trennlinie ist somit nicht immer eine Linie, sondern ein Polygonzug mit zwei bis vier Stützpunkten. Da die Verkürzung jedoch nur die Linienenden betrifft, bleibt die charakteristische Gestalt der Linie erhalten. Aus diesem Grund wie auch aus Aufwandsgründen werden die Trennlinienkan-

<sup>3</sup>In Abbildung 5.7 wird nur eine Segmenthälfte betrachtet. Tatsächlich reicht die Linie vom einem Segmentrandpixel  $B_1$  bis zum gegenüberliegenden  $B_2$ . Die Länge der Trennlinie,  $l_i$ , ist also äquivalent zu  $|\overline{B_1 B_2}|$ , bzw.  $|\overline{B_1 E}| + |\overline{EB_2}|$ .



**Abbildung 5.8:** Schematische Darstellung zum Ausschluss von Linien, die länger entlang des Segmentrands verlaufen. Durch zwei Parallelen (rote Linien) zur Trennlinie  $T_i$  im Abstand  $d^{(outer)}$  wird der segmentrandnahe Anteil ermittelt (rot gestrichelt).

didaten in den folgenden Abschnitten nur anhand des inneren Teils, also der eigentlichen Linie, beurteilt.

#### 5.4.3 Ausschluss von Linien nahe des Segmentrands

Die Trennlinienkandidaten werden im folgenden Abschnitt durch entsprechende Gütekriterien bewertet. Insbesondere spielt dabei die Gradientenstärke entlang der Linie eine Rolle. Bei Linien, die zu einem großen Teil entlang der Gebäudesegmentgrenze verlaufen (wie in Abbildung 5.8 dargestellt), wird dieses Gütemaß stark verzerrt. In der Nähe einer realen Gebäudekontur liegt wegen der dort vorhandenen Grauwertkante immer eine hohe Gradientenstärke vor. Diese hohe Gradientenstärke ist in solch einem Fall also kein Hinweis auf das Vorliegen einer Trennlinie.

Der Teil der Trennlinie, der sich nahe der Segmentgrenze befindet, wird über zwei Parallelen zur Trennlinie bestimmt. Dazu müssen die Parallelen einen relativ geringen Abstand  $d^{(outer)}$  zur Linie haben. Als „segmentrandnah“ wird der Teil der Linie bezeichnet, für den sich *mindestens eine der beiden Parallelen außerhalb des Gebäudesegments* befindet. Ist dieser Teil mit Länge  $l_i^{(outer)}$  größer als eine vorgegebene Schwelle  $l_{max}^{(outer)}$ , wird die Linie  $T_i$  verworfen.

### 5.5 Kriterien zur Beurteilung von Trennlinien

Zur Beurteilung der Güte der Trennlinienkandidaten werden im Folgenden drei Kriterien definiert. Das einfachste Kriterium ist die *Länge*  $l_i$  einer Linie. Auf das Längenkriterium wird zum Ende dieses Abschnitts nochmals eingegangen. Des Weiteren werden als Kriterium die *gemittelte Gradientenstärke* sowie der *mittlere Grauwertunterschied* hergeleitet.

#### 5.5.1 Mittelung der Gradientenstärke

Zur Bewertung der Trennlinien wird als ein Kriterium die Gradientenstärke entlang dieser Strecke herangezogen. Im letzten Abschnitt wurde die Gradientenstärke bereits als Selektionskriterium be-

nutzt. Die Mittelung der Gradientenstärke ist dabei jedoch aus Gründen einer schnellen Vorselektion innerhalb einer festen Nachbarschaft  $N^{(1)}$  um ein Basispixel erfolgt. Bei der jetzt stattfindenden Mittelung werden ausschließlich die von der Trennlinie betroffenen Pixel,  $T$ , einbezogen. Da das Bestimmen der von Trennlinie zu Trennlinie verschiedenen Linienpixel aufwändiger ist als die Verwendung einer festen Nachbarschaft, erfolgt die Selektion anhand der Gradientenstärke in zwei Einheiten. Für die linienspezifische Mittelung der Gradientenstärke ergibt sich die absolute Differenz von (der aus den Segmentrandpixeln bestimmten) Hauptorientierung zu der Orientierung eines Trennlinienpixels  $j \in T$  nach Gleichung 5.2. Die über die Trennlinienpixel *gemittelte Gradientenstärke* berechnet sich analog zu Gleichung 5.3:

$$|\overline{\nabla}|^{(T_i)} = \frac{1}{|T_i|} \sum_{j \in T_i} |\nabla|_j \left( 1 - \frac{\min(\Delta\varphi_{ij}, \Delta\varphi_{\max})}{\Delta\varphi_{\max}} \right). \quad (5.8)$$

Um Fehler in der Abschätzung der Linienausrichtung zu kompensieren, werden auch die Pixel seitlich der Trennlinie in den Mittelungsprozess einbezogen. Allerdings werden die zur Trennlinie benachbarten Pixel schwächer gewichtet als die Pixel der Trennlinie.

Die Konturen von Dachelementen, wie Kaminen oder Antennen, weisen ebenfalls eine hohe Gradientenstärke auf. Im ungünstigen Fall kann auch deren Gradientenorientierung vereinbar sein mit  $\hat{\varphi}$ . Dann resultiert eine hohe gemittelte Gradientenstärke. Um zu vermeiden, dass Trennlinien an Dachelementen konstruiert werden, wird ferner die *Verteilung der Pixel mit hoher Gradientenstärke und passender Gradientenorientierung* betrachtet. Dazu wird für einen Trennlinienkandidat  $T_i$  ein binärer Vektor  $T_i^{(g)}$  definiert. Für die in Linienabfolge geordneten Pixel der Trennlinie  $\{j \in T_i, <\}$  ergeben sich die  $n_i^{(g)} = |T_i^{(g)}|$  Elemente  $T_{ij}^{(g)}$  folgendermaßen:

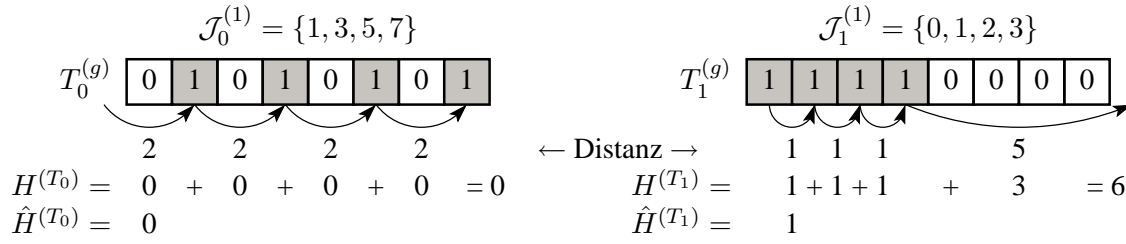
$$T_{ij}^{(g)} = \begin{cases} 1 & \text{falls } (|\nabla|_j > |\nabla|_{\min}) \wedge (\Delta\varphi_{ij} < \Delta\varphi_{\max}), \\ 0 & \text{sonst} \end{cases} \quad (5.9)$$

Die Indizes der passenden Kantenpixel (in  $T_i^{(g)}$  mit „1“ markiert) werden in der gemäß ihrem Betrag aufsteigend geordneten Menge  $\mathcal{J}_i^{(1)}$  zusammengefasst:

$$\mathcal{J}^{(T_i)} = \left\{ j \mid T_{ij}^{(g)} = 1 \right\} \quad (5.10)$$

Sind diese Pixelpositionen homogen in  $T_i^{(g)}$  verteilt, gilt dies als Indiz für eine durchgehende Trennlinie. Andernfalls wird davon ausgegangen, dass lokale Dachelemente zu dem hohen  $|\overline{\nabla}|^{(T_i)}$  geführt haben. Ein Maß für die Verteilung der passenden Kantenpixel in  $T_i^{(g)}$  ergibt sich über deren Abstand zueinander. Bei einer Anzahl von  $n_i^{(1)} = |T_i^{(1)}|$  Kantenpixeln sei der mittlere Abstand der zugehörigen Pixelindizes  $\bar{d}_i^{(1)}$ . Das Homogenitätsmaß  $H^{(T_i)}$  ergibt sich dann aus der Abweichung der realen





**Abbildung 5.9:** Veranschaulichung zur Beurteilung der Verteilung passender Kantenpixel (hier grau hinterlegt und mit „1“ markiert) entlang der Trennlinie. Der linke Fall gibt den bestmöglichen Fall, die Verteilung in gleichen Abständen, wieder. Der rechte Fall zeigt den schlechtesten Fall, die Ansammlung der Kantenpixel in einem Cluster. In beiden Fällen sind  $n_i^{(g)} = 4$  Kantenpixel bei insgesamt  $n_i^{(1)} = 8$  Linienelementen vorhanden. Der mittlere Abstand  $\bar{d}_i^{(1)}$  ist somit jeweils 2.

Abstände zu  $\bar{d}_i^{(1)}$ :

$$\bar{d}_i^{(1)} = \frac{n_i^{(g)}}{n_i^{(1)}} \quad (5.11)$$

$$H^{(T_i)} = \sum_{k \in \mathcal{J}^{(T_i)}} \left| \left( \left( \mathcal{J}_{k+1}^{(T_i)} - \mathcal{J}_k^{(T_i)} + n_i^{(g)} \right) \bmod n_i^{(g)} \right) - \bar{d}_i^{(1)} \right| \quad (5.12)$$

Das Maß  $H$  ist niedrig für gut über die Trennlinie verteilte Kantenpixel und hoch für lokal konzentriert auftretende Kantenpixel. Um die  $H^{(T_i)}$  als Wichtungsfaktoren für  $|\nabla|^{(T_i)}$  zu verwenden, werden sie auf das Intervall  $[0, 1]$  normiert:

$$\hat{H} = \begin{cases} \frac{H - H_{\min}}{H_{\max} - H_{\min}} & \text{falls } H_{\max} > H_{\min}, \\ 0 & \text{sonst} \end{cases} \quad (5.13)$$

Die Begrenzung nach unten,  $H_{\min}$ , kann durch die nur diskret vorkommenden Abstände der Pixelindizes auch größer als Null sein. Die Begrenzung nach oben,  $H_{\max}$ , ist im Fall nur eines Clusters mit passenden Kantenpixeln erreicht. Die beiden Grenzen (jeweils in Abhängigkeit von  $n_i^{(1)}$  und  $n_i^{(g)}$ ) ergeben sich wie folgt:

$$H_{\min}^{(T_i)} = n_i^{(g)} - n_i^{(1)} \left\lfloor \frac{n_i^{(g)}}{n_i^{(1)}} \right\rfloor \quad (5.14)$$

$$H_{\max}^{(T_i)} = \frac{2 \left( n_i^{(1)} - 1 \right) \left( n_i^{(g)} - n_i^{(1)} \right)}{n_i^{(1)}} \quad (5.15)$$

In Abbildung 5.9 ist die Ermittlung des Homogenitätsmaßes an zwei Beispielfällen mit jeweils gleicher Anzahl passender Kantenpixel veranschaulicht.

Das so gewonnene Homogenitätsmaß wird nun zur Gewichtung der mittleren Gradientenstärke

verwendet. Die gewichtete Gradientenstärke  $|\nabla|_H^{(T_i)}$  ist definiert als

$$|\nabla|_H^{(T_i)} = \begin{cases} |\nabla|^{(T_i)} & \text{falls } \bar{d}_i^{(1)} > \bar{d}_{\max}^{(1)}, \\ |\nabla|^{(T_i)} (2 - \hat{H}^{(T_i)}) & \text{sonst.} \end{cases} \quad (5.16)$$

### 5.5.2 Differenz der beidseitigen Grauwerte

Neben dem Vorhandensein einer Bildkante sind deutlich unterschiedliche Grauwerte zu beiden Seiten der Trennlinie ein Indiz für andere Dachmaterialien und somit für eine real existierende Trennlinie. So sind in dem in den Abbildungen 5.13 und 5.14 folgenden Bildbeispiel die hellen und dunklen Dächer sofort voneinander zu unterscheiden. Aus dem *Unterschied der beidseitigen Grauwerte* wird ein weiteres Kriterium zur Beurteilung der Trennlinien abgeleitet. Dazu werden die Grauwerte von den Pixeln betrachtet, die sich durch eine Projektion der Linienelemente  $T_{ij}$  senkrecht zur Trennlinie ergeben. Die Projektion eines Linienpixels an der Position  $\mathbf{x}_{ij} = (x_i, y_i)$  resultiert aus der Addition mit dem Verschiebungsvektor  $\mathbf{ds}$ :

$$\mathbf{ds} = (\cos(\hat{\varphi}_i + \pi/2), \sin(\hat{\varphi}_i + \pi/2)) \quad (5.17)$$

Die Grauwerte werden dabei in geringer Entfernung (3–5 Pixel) von der Trennlinie abgegriffen, um zu vermeiden, dass auch Kantenpixel der Trennlinie betrachtet werden. Damit möglichst nur Grauwerte innerhalb des Gebäudesegments einbezogen werden, werden zudem nur die inneren Linienpixel (Abstand von zwei Pixeln zur Segmentgrenze) projiziert. Aus den Mittelwerten von jeweils drei in Projektionsrichtung hintereinander liegenden Pixeln (s. Abb. 5.3) werden die beiderseitigen Grauwert-Vektoren  $\bar{f}^{(0)}$  und  $\bar{f}^{(1)}$  gebildet:

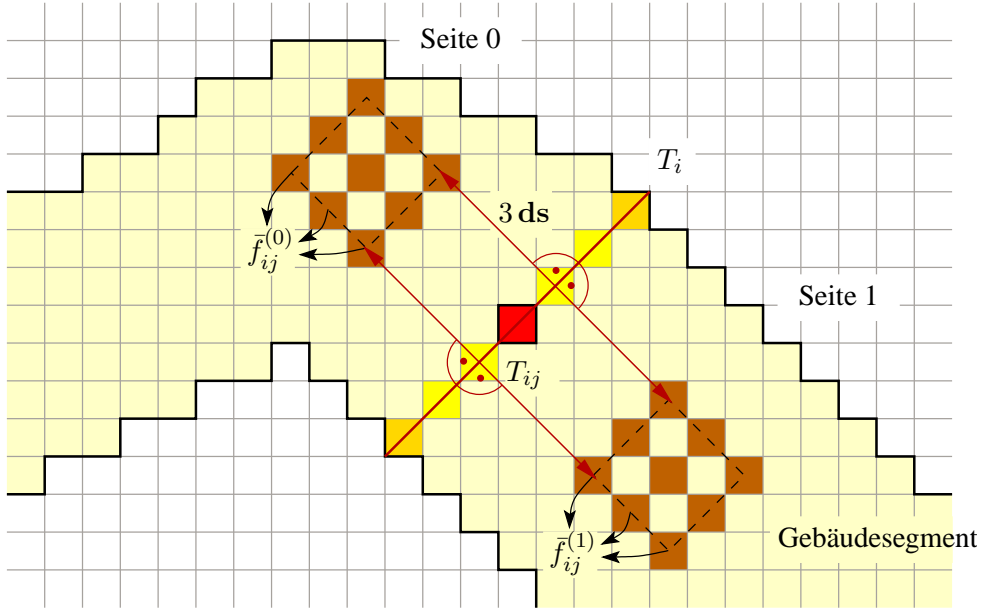
$$\bar{f}_{ij}^{(0)} = \frac{1}{3} \sum_{k=3}^5 f(\mathbf{x}_{ij} + k \mathbf{ds}), \quad \bar{f}_{ij}^{(1)} = \frac{1}{3} \sum_{k=3}^5 f(\mathbf{x}_{ij} - k \mathbf{ds}) \quad (5.18)$$

Aus diesen beiden Grauwert-Vektoren resultiert die mittlere DN-Differenz  $\Delta \bar{f}$ , die als Kriterium zur Beurteilung der Trennlinien herangezogen wird:

$$\Delta \bar{f}_i = \sum_j \left| \bar{f}_{ij}^{(0)} - \bar{f}_{ij}^{(1)} \right| \quad (5.19)$$

Je höher dieses Maß ausfällt, desto eher liegen unterschiedliche Dachmaterialien vor und somit auch ein Übergang von einem Haus zu einem anderen. Das Vorgehen zur Ermittlung der beidseitigen Grauwertdifferenz ist an einem Bildbeispiel in Abbildung 5.10 illustriert.

Zusammenfassend stehen nun drei Kriterien zur Bewertung der Trennlinien zur Verfügung. Zum einen die *gemittelte Gradientenstärke*, die durch die Gradientenorientierung und die Homogenität passender Linienpixel gewichtet wird, der *mittlere Grauwertunterschied* zu beiden Seiten der Trenn-



**Abbildung 5.10:** Schematische Darstellung zur Bildung der beidseitigen Grauwertdifferenz  $\Delta \bar{f}_i$  einer Trennlinie  $T_i$ . Die braun hinterlegten Pixel werden zur Ermittlung der Grauwertdifferenz herangezogen.

linie sowie die *Länge* der Linie. Das Längenkriterium wurde bereits zum Ausschluss zu kleiner Linien verwendet (letzter Schritt gemäß Abschnitt 5.3). Hier wird es aufgenommen, um längere Linien kürzeren mit vergleichbarer Güte vorzuziehen. Denn je länger eine Linie ist, desto verlässlicher ist die Mittelung der Gradientenstärke. Wie bereits erwähnt, können Dachelemente wie Kamine oder Antennen im ungünstigen Fall auch zu einer hohen gemittelten Gradientenstärke führen. Je länger ein Trennlinienkandidat ist, desto unwahrscheinlicher ist es jedoch, dass er lediglich bedingt durch Dachelemente zustande gekommen ist.

## 5.6 Bewertung und Auswahl geeigneter Trennlinien

Durch die drei Gütekriterien, die im letzten Abschnitt vorgestellt wurden, werden die Trennlinienkandidaten nun bewertet. Dazu werden die Gütemaße durch Linearkombination zu einem Kriterium  $q^{(T)}$  vereinigt. Als abkürzende Schreibweise wird zuvor der Operator  $\uparrow \cdot \downarrow_{\theta_{\min}}^{\theta_{\max}}$  eingeführt:

$$\uparrow x \downarrow_{\theta_{\min}}^{\theta_{\max}} = \min(\max(x, \theta_{\min}), \theta_{\max}) \quad (5.20)$$

Falls eine der Schranken nicht spezifiziert ist, ist das Maß in diese Richtung unbeschränkt.

Das kombinierte Gütekriterium ergibt sich dann durch Normierung der einzelnen Kriterien und einer anschließenden Streckung auf einen für die jeweiligen Kriterien spezifischen Wertebereich

durch Gewichtungsfaktoren  $w$ :

$$q^{(T_i)} = \frac{\frac{\uparrow \downarrow |\nabla|_H^{(T_i)} \uparrow \downarrow |\nabla|_H^{\max}}{|\nabla|_H^{\max}}}{|\nabla|_H^{\max}} + \frac{\frac{\uparrow \downarrow \Delta \bar{f}_i \uparrow \downarrow \Delta \bar{f}_{\max}}{\Delta \bar{f}_{\max} - \Delta \bar{f}_{\min}}}{\Delta \bar{f}_{\max} - \Delta \bar{f}_{\min}} \left( w_{\max}^{\Delta \bar{f}} - w_{\min}^{\Delta \bar{f}} \right) + w_{\min}^{\Delta \bar{f}} + \frac{\frac{\uparrow \downarrow l_i \uparrow \downarrow l_{\max} - l_{\min}}{l_{\max} - l_{\min}}}{l_{\max} - l_{\min}} \left( w_{\max}^l - w_{\min}^l \right) + w_{\min}^l \quad (5.21)$$

Die Ober- und Untergrenzen für das jeweilige Kriterium resultieren aus den Erfahrungen in zahlreichen Testläufen. Die Gewichtungsfaktoren zur Kombination der Kriterien werden anhand eines Referenzdatensatzes ermittelt. Die Faktoren, die zu den höchsten Erkennungsraten bei möglichst geringen Fehlalarmraten führen<sup>4</sup>, werden als Standardwerte gewählt.

### 5.6.1 Entfernen benachbarter Linien

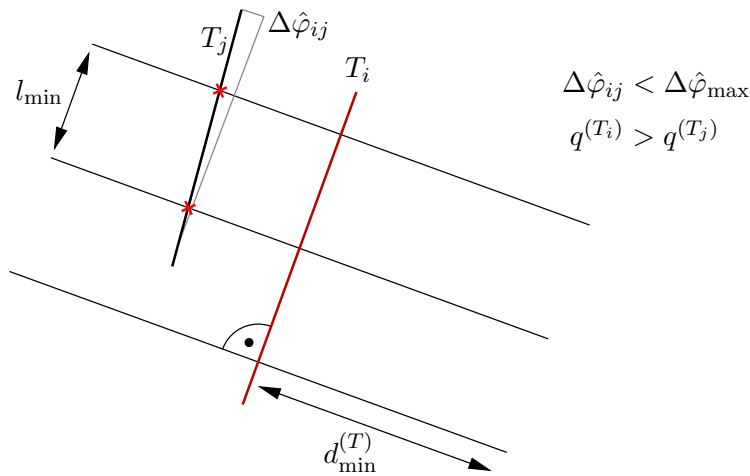
Durch das Kriterium  $q^{(T)}$  sind die Trennlinien entsprechend ihrer Güte geordnet. Wie in Abbildung 5.13(b) zu sehen ist, entstehen in der Nachbarschaft von realen Hausgrenzen meist zahlreiche Trennlinienkandidaten. Anhand des Gütekriteriums können die nicht-maximalen Linien einfach unterdrückt werden. Beginnend bei der bestbewerteten Linie werden in einer Nachbarschaft um diese Linie alle weiteren Linien mit ähnlicher Orientierung (Winkelunterschied  $< \Delta \hat{\varphi}_{\max}$ ) gelöscht. Dazu wird geprüft, wie in Abbildung 5.11 veranschaulicht, ob Schnittpunkte von Linien ähnlicher Orientierung mit Senkrechten zur Trennlinie  $T_i$  existieren. Die Senkrechten haben den Abstand  $l_{\min}$ , so dass auch die kürzest möglichen Linien noch detektiert und entfernt werden können<sup>5</sup>. Die Senkrechten haben eine Länge  $2d_{\min}^{(T)}$ , womit garantiert wird, dass die nächste Trennlinie mindestens im Abstand  $d_{\min}^{(T)}$  von der aktuellen Linie ist.

### 5.6.2 Entfernen von Linien, die zu zu kleinen Segmenten führen würden

Segmente, die kleinere Abmessungen aufweisen als ein einzelnes Haus, sollten bei der Unterteilung der Gebäudekomplexe nicht auftreten. Durch die gerade vorgestellte Unterdrückung nicht-maximal bewerteter Trennlinien werden bereits zu kleine Segmente — hervorgerufen durch benachbarte Trennlinienkandidaten — ausgeschlossen. Nun soll auch die Größe äußerer Hauseinheiten betrachtet werden, also Segmente, die nur durch eine Trennlinie und sonst den Segmenträndern begrenzt sind. Diesbezüglich wird bereits in Abschnitt 5.4 der Ausschluss von Linien ohne beidseitige Segmentfortsetzung diskutiert. Der Ausschluss zu kleiner Segmente erfolgt analog zu dem dort wie auch im vorhergehenden Abschnitt genannten Vorgehen. Statt einer Senkrechtenschar kommt hier lediglich die Mittelsenkrechte der jeweiligen Trennlinie zum Einsatz. Im Abstand  $d_{\min}^{(T)}$  beidseitig

<sup>4</sup>Auf die Bestimmung der Erkennungsraten bei Trennlinien wird in Abschnitt 5.8 eingegangen.

<sup>5</sup>Korrektweise darf der Abstand zur Detektion der kleinstmöglichen Linien bei einer Winkeldifferenz  $\Delta \hat{\varphi}_{ij}$  höchstens  $l_{\min} / \cos(\Delta \hat{\varphi}_{ij})$  sein. Da bei kleinen Winkelunterschieden jedoch  $\cos(\Delta \hat{\varphi}_{ij}) \approx 1$  ist, wird der Einfachheit halber  $l_{\min}$  als Abstandsmaß gewählt.



**Abbildung 5.11:** Schematische Darstellung zur Bestimmung benachbarter Linien ähnlicher Orientierung. Basierend auf der aktuellen Trennlinie  $T_i$  werden Senkrechten der Länge  $2d_{\min}^{(T)}$  im Abstand  $l_{\min}$  aufgespannt. Falls es zu Schnittpunkten mit anderen Trennlinien ähnlicher Orientierung, hier  $T_j$  kommen sollte, werden diese gelöscht.

der Linie wird überprüft, ob in einer kleinen Nachbarschaft ein Pixel vorhanden ist, das nicht mehr zum Gebäudesegment gehört. Ist dies der Fall, wird diese Linie gelöscht. Da der Abstand  $d_{\min}^{(T)}$  größer ist als der in Abschnitt 5.4 verwendete Mindestabstand  $d^{(\text{inner})}$ , ist diese abermalige Überprüfung notwendig.

Aufgrund des damit verbundenen Rechenaufwands wird auf eine explizite Ermittlung der Segmentgröße verzichtet. Alleine über die beiden in diesem Abschnitt vorgestellten Abstandskriterien wird die Mindestgröße eines Haussegments kontrolliert. Diese Überprüfungen sind effizient durch Projektionen und Schnittpunktbestimmung zu realisieren.

## 5.7 Realisierung der Trennlinien

Die in den vorangegangenen Schritten auf wenige Kandidaten reduzierten Trennlinien sind zwar gut bewertet, doch es sind noch zwei Aspekte abzuklären: Gibt es Schnitte mit anderen Trennlinien? In Realität ist dieser Fall möglich, aber relativ selten. In diesem Abschnitt soll dies der Einfachheit halber ausgeschlossen werden. Ein weiterer Aspekt: Beinhaltet der Gebäudekomplex auch Innenhöfe, so sind diese nicht Teil des Gebäudesegments. Reicht nun eine Trennlinie von der Außenseite des Gebäudes zu einem Innenhof (s. z. B. Linie 2 in Abbildung 5.14) so ist diese Linie für sich alleine genommen nach Definition 5.1 keine Trennlinie. Erst die Kombination mit weiteren Linien, die aus dem Innenhof wieder hinaus führen, führt zu einer echten Unterteilung des Gebäudesegments.

Um diesen beiden Aspekten Rechnung zu tragen, wird über die Akzeptanz eines Kandidaten schrittweise in Reihenfolge der Bewertung entschieden. Zur Unterscheidung, welche Segmente ei-

ne Linie verbindet, werden die Nicht-Gebäudesegmente numeriert. Dann wird jede Linie mit dem Attribut versehen, welche Segmente sie verbindet. Die Funktion **CONNECT** dient der Abfrage dieses Attributs:  $\text{CONNECT}(T_i) \mapsto \{s, t\}$ , wobei  $T_i$  die Segmente  $\text{seg}_s$  und  $\text{seg}_t$  verbindet. Die äußere Umgebung des Gebäudesegments wird standardmäßig mit  $\text{seg}_0$  bezeichnet, die Innenhöfe mit  $\text{seg}_u$ ,  $u > 0$  (s. Bildbeispiel in Abbildung 5.12).

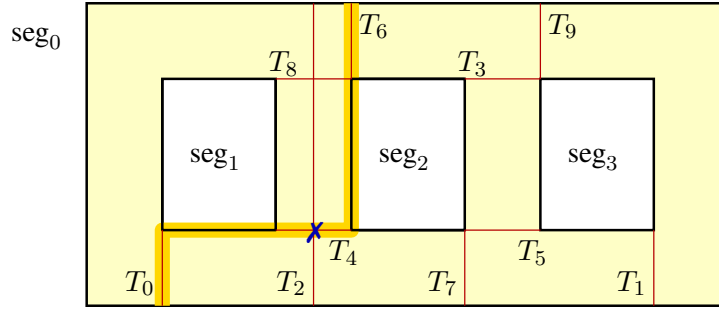
**Eingabe:** Trennlinien  $T = \{T_1, T_2, \dots, T_n\}$  nach Güte geordnet  
**REALISIERTRENNLINIEN**( $T$ )

```

(1)   while  $T \neq \emptyset$ 
(2)       switch  $\text{CONNECT}(T_0)$ 
(3)            $\{0, 0\}$ :
(4)                $R \leftarrow \{T_0\}$ 
(5)                $D \leftarrow \{T_0\}$ 
(6)            $\{s, t\}$ , mit  $s, t > 0$ :
(7)               if  $\{T_i \mid \text{CONNECT}(T_i) = \{0, u\}\} \neq \emptyset$  mit  $u > 0$ 
(8)                   Tausche  $T_0$  mit nächster Linie  $T_j$  für die gilt  $\text{CONNECT}(T_j) =$ 
                         $\{0, v\}$  mit  $v \geq 0$ 
(9)               continue
(10)           $D \leftarrow \{T_i \mid \text{CONNECT}(T_i) = \{u, v\}\}$ , mit  $u, v > 0$ 
(11)           $\{0, s\}$ , mit  $s > 0$ :
(12)              Bestimme besten Pfad  $P_{\text{best}}$  beginnend mit  $T_0$ 
(13)              if  $P_{\text{best}} \neq \emptyset$ 
(14)                   $R \leftarrow P_{\text{best}}$ 
(15)                   $D \leftarrow P_{\text{best}}$ 
(16)              else
(17)                   $D \leftarrow \{T_0\}$ 
(18)          if  $R \neq \emptyset$ 
(19)               $D \leftarrow D \cup \{T_i \mid (T_i \in T \setminus R) \wedge (T_i \vdash T_j)\}$ , mit  $T_j \in R$ 
(20)              Realisiere Linien  $R$ 
(21)               $\forall T_k \in T \setminus R$ : Aktualisieren der CONNECT-Attribute
(22)              Lösche Linien  $D$ 
```

**Algorithmus 5:** Vorgehen zur Realisierung der Trennlinienkandidaten.

Die Realisierung der Trennlinien, also die endgültige Unterteilung des Gebäudesegments, erfolgt in der Reihenfolge ihrer Bewertung nach Algorithmus 5. Dabei wird unterschieden, ob eine Verbindung von Außen- zu Außensegment (Zeile 3–5), von Innen- zu Innensegment (Zeile 6–10) oder von Außen- zu Innensegment (Zeile 11–17) vorliegt. Im ersten Fall wird die Linie sofort umgesetzt. Im zweiten Fall wird keine Linie realisiert; falls noch Verbindungen nach außen vorhanden sind, wird die aktuelle Linie zurückgestellt. Anderfalls werden alle inneren Verbindungen gelöscht. Im dritten Fall wird überprüft, ob ein Pfad von Trennlinien durch das Gebäudesegment möglich ist. Falls ja, werden alle am bestbewerteten Pfad beteiligten Linien umgesetzt. Abschließend werden Linien, die



**Abbildung 5.12:** Schematische Darstellung zur Ermittlung des besten Pfades zur Kombination „innerer“ Trennlinien. Ausgehend von der aktuellen Trennlinie  $T_0$  (die Linien sind nach ihrer Güte geordnet) wird ein Pfad von der Umgebung  $\text{seg}_0$  durch das Gebäude mit den Innensegmenten  $\text{seg}_1$ – $\text{seg}_3$  wiederum nach außen gesucht. Der Pfad entlang der Linien  $T_0$ ,  $T_4$  und  $T_6$  ist der Pfad mit den wenigsten daran beteiligten Linien und der besten Güte. Die Linie  $T_2$  kreuzt diesen Pfad und wird daher gelöscht.

die umzusetzenden Linien kreuzen (symbolisiert durch „ $\vdash$ “), gelöscht (Zeile 19).

Die Ermittlung eines Trennlinienpfades durch das Gebäudesegment beginnt also immer mit einer Verbindung von außen in ein inneres Segment. Der Pfad mit der geringsten Anzahl beteiligter Linien wird immer präferiert. Falls aus dem ersten inneren Segment wieder eine Linie zum umgebenden Außensegment führt, ist die Suche somit abgeschlossen. Ist dies nicht der Fall, wird überprüft, ob es eine Verbindungslinie zu einem weiteren inneren Segment gibt. Über ein als Breitensuche implementiertes Vorgehen wird rekursiv überprüft, ob eine Verbindung zum äußeren Segment möglich ist. Abbildung 5.12 veranschaulicht das Vorgehen an einem Beispiel.

Bei mehreren Pfaden mit gleicher Anzahl an Linien wird der bestbewertete Pfad gewählt. Die Bewertung erfolgt analog zu der Bewertung einzelner Trennlinien nach Gleichung 5.21. Die jeweiligen Qualitätskriterien werden dazu wie folgt zu einem Kriterium für den gesamten Pfad kombiniert:

$$l_{\text{Path}_i} = \sum_{T_j \in \text{Path}_i} l_j \quad (5.22)$$

$$\Delta \bar{f}_{\text{Path}_i} = \frac{1}{l_{\text{Path}_i}} \sum_{T_j \in \text{Path}_i} \Delta \bar{f}_j l_j \quad (5.23)$$

$$\overline{|\nabla|}_H^{(\text{Path}_i)} = \frac{1}{l_{\text{Path}_i}} \sum_{T_j \in \text{Path}_i} \overline{|\nabla|}_H^{(T_j)} l_j \quad (5.24)$$

Die Anwendung der in diesem Abschnitt beschriebenen Vorgehensweise an realen Bilddaten folgt im nächsten Abschnitt.

## 5.8 Evaluation der Gebäudeunterteilung

Die Entwicklung von Trennlinienkandidaten und deren schrittweise Selektion wird in Abschnitt 5.8.1 an einem Bildbeispiel illustriert. Dann folgen Angaben zur Genauigkeit (Abschnitt 5.8.2) und Rechenzeit (Abschnitt 5.8.3) der vorgestellten Algorithmen.

### 5.8.1 Qualitative Beurteilung

Zur Demonstration der Vorgehensweise dient der in den Abbildungen 5.13 und 5.14 dargestellte manuell segmentierte Gebäudekomplex. Die verwendeten Nadirdaten haben 0,5 m Pixelauflösung und zeigen einen typischen Berliner Gebäudeblock. Dieser Block ist durch die ihn umgebenden spitzwinklig verlaufenden Straßen nicht vollständig rechtwinklig aufgebaut. Benachbarte Häuser des Blocks heben sich teilweise gut durch ihre Dachmaterialien voneinander ab, teilweise aber auch fast gar nicht. Die Brandmauern sind relativ gut zu erkennen.

In Abbildung 5.13(a) sind die Skelettpixel den Bilddaten farblich überlagert. Rot dargestellt sind die Skelettpixel, die entsprechend Abschnitt 5.3 zwei Kriterien erfüllen. Zum einen verträgt sich ihre Gradientenorientierung mit der naher Segmentrandpixel. Zum anderen liegt die in einer kleinen Nachbarschaft gemittelte Gradientenstärke  $|\nabla|^{N_i^{(1)}}$  über einem Mindestmaß. Von den ursprünglichen 1390 Skelettpixeln verbleiben dadurch 747 als Kandidaten für eine Trennlinie. Für diese wird eine entsprechender Polygonzug generiert. Nach Ausschluss der zu kurzen Linien wird die Ausrichtung der Linien vereinheitlicht. Nach den in den Abschnitten 5.5 und 5.6 vorgestellten Verfahren wird für jede Trennlinie das Gütekriterium  $q^{(T_i)}$  berechnet. Die 171 Linien mit hinreichender Güte sind in Abbildung 5.13(b) dargestellt. Beginnend bei den bestbewerteten Linien werden benachbarte Linien ähnlicher Orientierung entfernt; die verbleibenden 34 Trennlinien sind in Abbildung 5.13(c) zu sehen. Schließlich werden noch die Linien, die zu zu kleinen Segmenten führen würden (Linien 1, 12, 26, und 28 in Abb. 5.13(c)), entfernt, s. Abb. 5.13(d). Nun erfolgt die in Abschnitt 5.7 beschriebene schrittweise Umsetzung von Liniengruppen. Das endgültige Ergebnis der Unterteilung des Gebäudeblocks zeigt Abbildung 5.14.

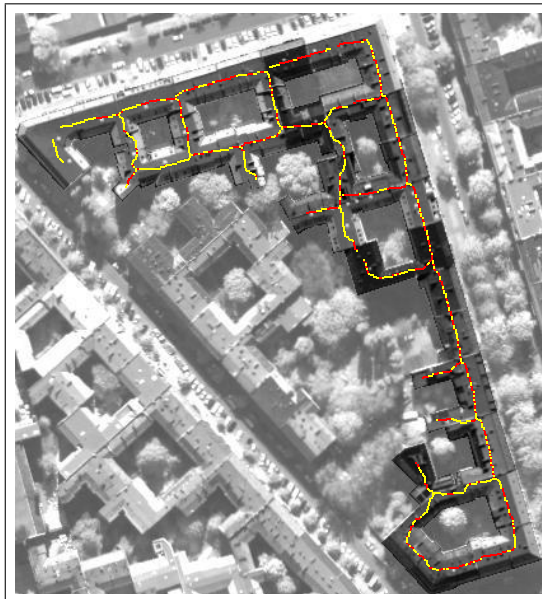
Die wesentlichen Hausgrenzen sind in diesem Beispiel erfasst worden. Die links neben Linie 22 in Abbildung 5.14 liegende Gebäudeunterteilung wurde nicht detektiert, da sie zu kontrastarm und auch visuell nur schwer auszumachen ist. Linie 20 wäre etwas weiter links oder rechts besser positioniert gewesen. Für Linie 15 ist die Ausrichtung nicht ganz passend. Für sie wie auch für die Linien 1 und 6 ist die Linienendenverkürzung zu früh erfolgt, bzw. wäre nicht notwendig gewesen<sup>6</sup>.

Dass die drei längeren Linien 4, 15 und 28 in Abbildung 5.13(d) umgesetzt werden und nicht die sie kreuzenden inneren Linien 7 und 8, liegt daran, dass die Verbindungen zur äußeren Gebäudeumgebung nur durch die schlechter bewerteten Linien 17, 19, 23, 24, 25 und 26 möglich ist (die Linien

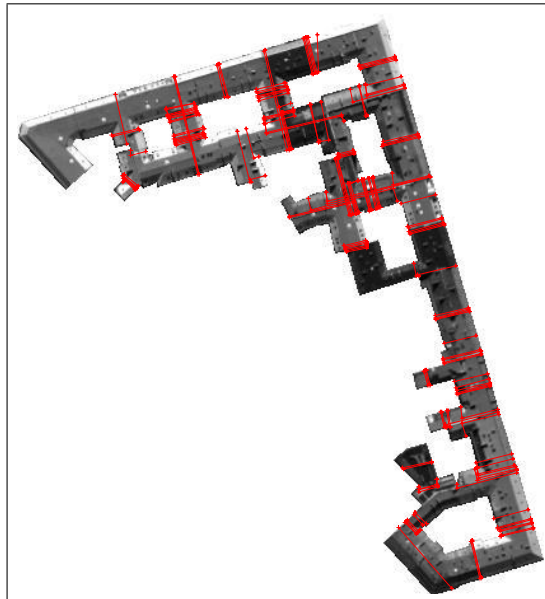
---

<sup>6</sup>Linie 1 kann jedoch nur so zustande kommen. Andererseits würde sie im Inneren des Segments enden, was nicht Definition 5.1 entspräche.

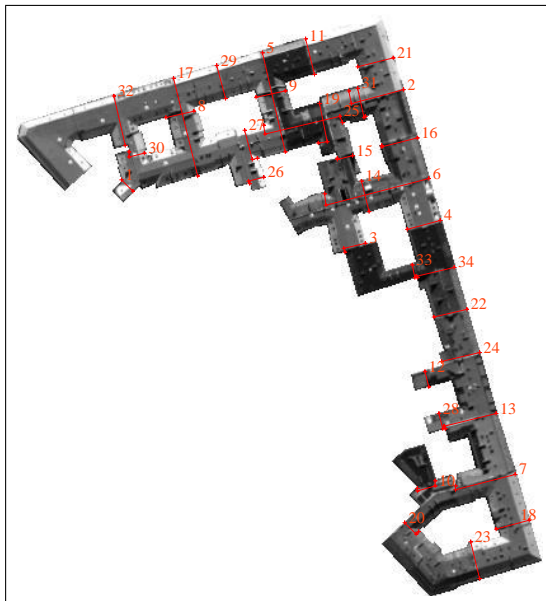




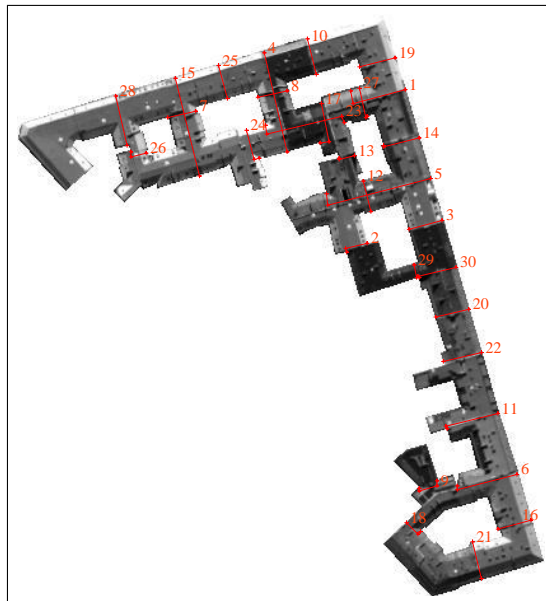
(a) Gebäudesegment und Skelettpixel



(b) Linien ausreichender Länge und Qualität

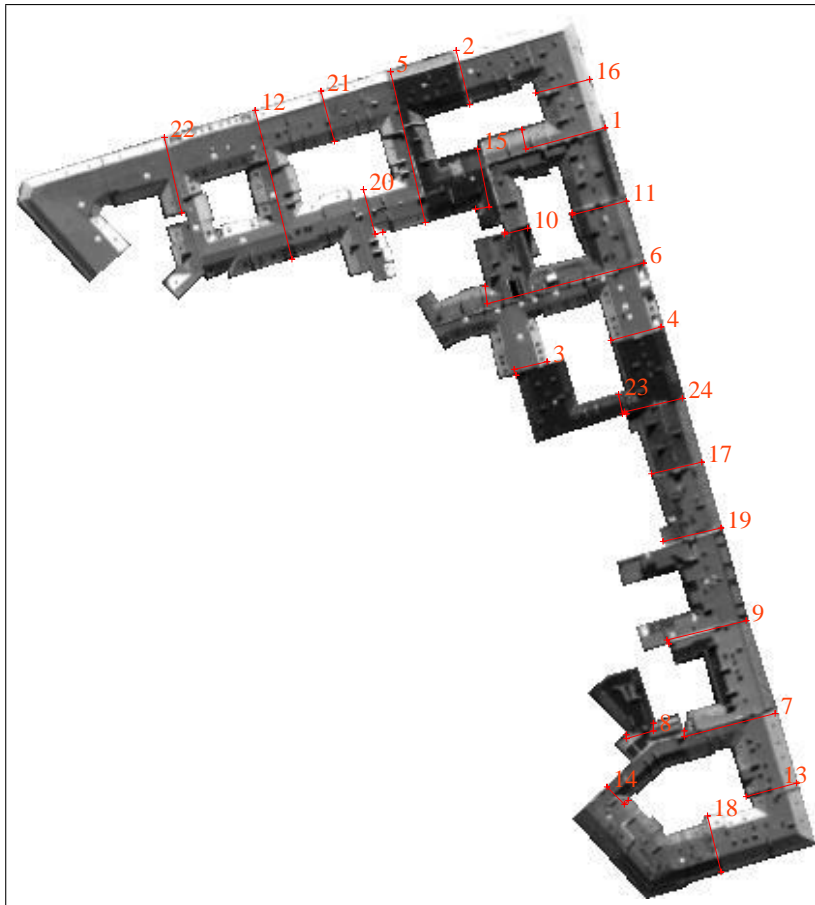


(c) Entfernung von Linien ähnlicher Orientierung



(d) Ausschluss zu kleiner Segmente

**Abbildung 5.13:** Zwischenresultate bei der Unterteilung eines Gebäudekomplexes in kleinere Einheiten. Dem Gebäudesegment überlagert sind in (a) die Skelettpixel (gelb und rot). Daraus werden die rot markierten Skelettpixel ausgewählt. In (c)–(d) folgen weitere Selektionsschritte auf Trennlinienebene. In den Abbildungen (c) und (d) sind die Linien entsprechend ihres Qualitätskriteriums numeriert.

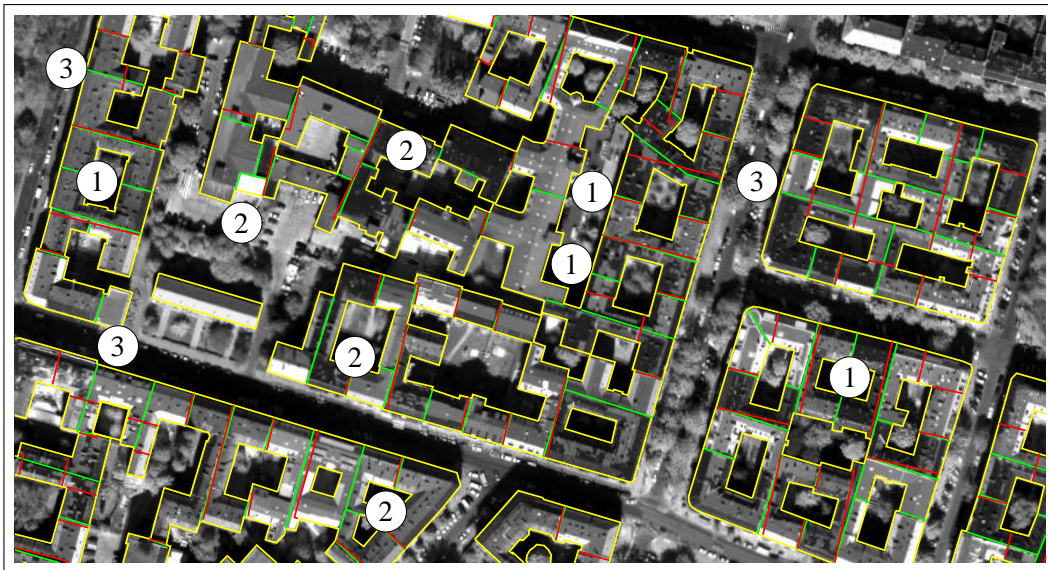


**Abbildung 5.14:** Endergebnis bei der Unterteilung eines Gebäudekomplexes in kleinere Einheiten. Entsprechend Abschnitt 5.7 werden nur konsistente und sich nicht kreuzende Liniengruppen realisiert (Reihenfolge der Numerierung entspricht dem Qualitätskriterium).

1 und 10 wurden bereits realisiert).

Abbildung 5.15 zeigt einen Ausschnitt des nachfolgend in Abschnitt 5.8.2 eingeführten Referenzdatensatzes. Die gelb konturierten Gebäudesegmente sind manuell extrahiert. Überlagert sind sowohl die automatisch bestimmten Trennlinien wie auch die Vorgaben. Neben der Übereinstimmung von Vorgabe und Unterteilungsergebnis sind drei typische Fehlersituationen auszumachen:

1. Die Linienvorgabe wird nicht detektiert.
2. Die Linienvorgabe endet im Inneren des Gebäudesegments. Da für das vorgeschlagene Vorgehen vorausgesetzt wird, dass die Trennung eine Verbindung von Außenbereich und/oder Innenhöfen darstellt, können diese Linien nicht erkannt werden.
3. Bei dem vorgestellten Vorgehen werden alle Linien ausgeschlossen, die bereits realisierte Linien kreuzen.



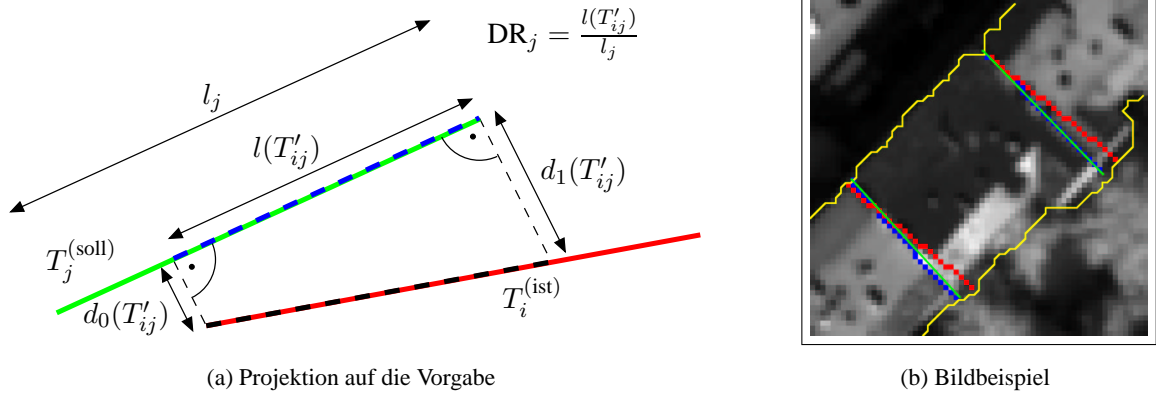
**Abbildung 5.15:** Auswertung der Unterteilung von Gebäudekomplexen an einem Referenzdatensatz (gezeigt ist ein Ausschnitt davon). Die Gebäudesegmentierung wurde hier manuell erzeugt (die Gebäudegrenzen sind gelb dargestellt). Die von Hand vorgegebenen Trennlinien sind grün und die automatisiert ermittelten Linien rot dargestellt. Anmerkungen zu den mit einer Ziffer versehenen Sonderfällen sind dem Text zu entnehmen.

Diese Fälle sind in Abbildung 5.15 mit der entsprechenden Ziffer hervorgehoben.

### 5.8.2 Genauigkeit

Bei der nun folgenden Bewertung der Genauigkeit werden die vorhergehend genannten Fehlersituationen nicht unterschieden. Eine nicht abgedeckte Linienvorgabe führt zu geringeren Detektionsraten. Bei der Bewertung der Detektionsraten muss dieser Aspekt jedoch berücksichtigt werden. So ist es mit dem vorgestellten Verfahren überhaupt nicht möglich, bei realen Daten Detektionsraten nahe 100 % zu erzielen. Eine Auswertung, die die Beschränkungen des vorgestellten Ansatzes berücksichtigt, ist nur schwer möglich. So müsste bei sich kreuzenden Linien festgelegt werden, welche die wichtigere wäre. Das Entfernen von Linien bedingt zudem meist auch den Ausschluss eigentlich akzeptabler Linien, um die Konsistenz zu wahren. Aus diesen Gründen werden die Linienvorgaben nicht „bereinigt“. Für reale Anwendungen geben die resultierenden Genauigkeiten zudem eine realistischere Abschätzung wieder.

Da sowohl das manuelle wie auch das automatische Bestimmen von Trennlinien mit Ungenauigkeiten behaftet ist, müssen bei der Auswertung Toleranzschwellen definiert werden. Bis zu welcher Entfernung von der Vorgabe und welchem Winkelunterschied soll eine Linie als detektiert gelten? Was ist, wenn die detektierte Linie nur einen Teil der Vorgabe darstellt? Im Folgenden werden diese Fragen geklärt; anschließend werden die so definierten Detektionsraten an einem Referenzdatensatz



**Abbildung 5.16:** Veranschaulichung des Vorgehens zur Bewertung des Trennlinienalgorithmus. Abbildung (a) gibt die Projektion detektierte Trennlinien auf Linienvorgaben und umgekehrt schematisch wieder; in Abbildung (b) beispielhaft ein Resultat anhand realer Bilddaten (grün dargestellt ist die Linienvorgabe, rot der relevante Teil einer ähnlichen detektierten Linie, blau die Projektion davon auf die Vorgabe und gelb die extrahierte Kontur des Gebäudesegments).

ermittelt.

Die Ermittlung der Detektionsrate erfolgt zunächst für jede Linienvorgabe separat. Dazu dient die Abdeckung der Vorgabe durch ähnliche benachbarte Linien. Eine durch die detektierten Linien nur halb abgedeckte Linie führt also zu einer Detektionsrate von 50 % für diese Linie. Zur Berechnung der Abdeckung werden für jede Linienvorgabe  $T_j^{(soll)}$  die detektierten Linien  $T_i^{(ist)}$  bestimmt, deren Orientierung um weniger als  $20^\circ$  von der Vorgabe abweicht. Diese werden orthogonal auf die Vorgabe projiziert, wie in Abbildung 5.16(a) veranschaulicht. Die resultierenden Projektionen seien

$$T'_{ij} = \text{Proj}_{T_j^{(soll)}} T_i^{(ist)} . \quad (5.25)$$

Nur die Projektionen, die die Vorgabe auch abdecken, werden zur Bewertung dieser Linie herangezogen. Außerdem muss der mittlere Abstand

$$\bar{d}(T'_{ij}) = 1/2 (d_0(T'_{ij}) + d_1(T'_{ij})) \quad (5.26)$$

kleiner als 3,5 m sein. Aus den verbleibenden projizierten Linien (im Regelfall eine) ergibt sich der Grad der Abdeckung der Vorgabe. Eine anhand der Länge der Linienvorgaben gewichtete Mittelung der Liniendetektionsraten führt dann zur Gesamtdetektionsrate.

Die Fehlalarmrate errechnet sich analog. Für alle detektierten Linien wird der Anteil bestimmt, der nicht zur Ermittlung der Detektionsrate herangezogen wurde. Eine anhand der Länge der detektierten Linien gewichtete Mittelung dieser Anteile ergibt dann die Fehlalarmrate.

Zur Ermittlung der Genauigkeiten wird ein Referenzdatensatz mit möglichen Gebäudetrennlinien

manuell erstellt (s. Abbildung 5.15). Grundlage dafür ist auch die bereits im letzten Kapitel verwendete Stadtszene mit 0,5 m Pixelauflösung. Die Trennlinien werden in eine bereits vorliegende manuell ermittelte Gebäudesegmentierung eingepasst. Die Auswertung erfolgt einmal anhand der gerade erwähnten manuell gewonnenen Gebäudesegmente. Ein weiteres Mal wird eine automatisiert ermittelte Segmentierung herangezogen. Die anhand der manuellen Segmentierung bestimmten Referenz-Trennlinien passen bei den automatisch ermittelten Gebäudesegmenten nicht genau. Deswegen werden für diesen Fall die Längen der Referenzlinien an die Ausdehnung der automatisiert bestimmten Segmente angepasst und nicht mehr sinnvolle Linien entfernt.

Tabelle 5.1 gibt die Genauigkeitswerte für diesen Referenzdatensatz an. Dazu wurde unter Variation von mehreren Parametern und Schwellwerten die Konstellation gewählt, die die Summe aus Detektionsrate und  $(1 - \text{FAR})$  maximiert. In den Testläufen werden auch höhere Detektionsraten erreicht, dann allerdings zu Lasten der Fehlalarmrate.

Zugrunde liegende Gebäudesegmentierung	DR (%)	FRR (%)	FAR (%)
manuell	61,97	38,03	21,09
automatisch	57,39	42,61	28,55

**Tabelle 5.1:** Genauigkeitsangaben für die Unterteilung der Gebäudesegmente.

Grundlage für die Parameterbestimmung ist eine manuelle Segmentierung der Gebäude. Die so ermittelte Parameterkonstellation wird auch an einer nach Kapitel 4 automatisiert bestimmten Gebäudesegmentierung angewendet.

Für die manuelle Gebäudesegmentierung besteht die Trennlinienvorgabe aus 270 Linien, bei der automatisiert ermittelten Segmentierung aus 255 Linien. Die daran ermittelten Detektionsraten bewegen sich um 60 % bei Fehlalarmraten von etwas über 20 %. Die Genauigkeiten bei Verwendung der nach Kapitel 4 automatisiert ermittelten Gebäudesegmentierung sind etwas schlechter. Dies ist zu erwarten, denn die Form des Gebäudesegments hat auch einen zwar geringeren, aber dennoch nicht zu vernachlässigenden Einfluss auf die Generierung der Trennlinien. Die Position der Skelettpixel ist nicht mehr dieselbe und bezüglich der realen Gebäudeform also nicht mehr optimal. Weiterhin werden am Segmentrand nach Abschnitt 5.2 Regionen hoher Gradientenstärke berechnet. Bei einer zu sehr von der Realität abweichenden Gebäudesegmentierung können hierbei falsche Kanten erfasst oder aber vorhandene Gebäudekanten nicht erfasst werden.

Die Detektionsraten sind für eine vollautomatische Ermittlung einzelner Häuser noch nicht ausreichend. Für die im nächsten Kapitel folgende Verbesserung der Gebäudekontur ist ausschlaggebend, dass die Konturen möglichst einfach sind, d. h. durch wenige Polygonstützpunkte beschrieben werden können. Dieses Ziel ist mit der Untergliederung der Gebäudesegmente in kleinere Einheiten erreicht. Für die Aufnahme in ein Geoinformationssystem ist jedoch entscheidend, dass die Unter-

gliederung nur an Hausgrenzen erfolgt. Um dies zu gewährleisten ist eine manuelle Interaktion noch unabdingbar. Dafür kann der vorgestellte Ansatz ein Hilfsmittel sein. Ein Vorschlag zur Unterteilung der Gebäude kann so automatisch erstellt und nachfolgend manuell überarbeitet werden.

### 5.8.3 Rechenzeit

Für die einzelnen in diesem Kapitel vorgestellten Arbeitsschritte zur Unterteilung eines Gebäudesegments werden nun die Laufzeiten an einem Beispielfall gemessen. Grundlage ist ein Gebäudesegment von 40 676 Pixeln, also etwa 10 000 m<sup>2</sup>. Die Laufzeitmessung erfolgt mit derselben Parameterkonstellation wie die Bestimmung der Genauigkeit. Da die Laufzeiten nicht nur in Abhängigkeit der Segmentgröße, sondern auch (in geringerem Maße) in Abhängigkeit der gewählten Parameter variieren, können diese Angaben nur als Anhaltspunkt dienen.

Arbeitsschritt	Einheiten (Skelett-pixel/Linien)	Gesamtzeit (ms)	Mittlere Zeit pro Einheit (ms)
Skelettieren	(40 676)	738,1	
Selektion auf Skelettpixelebene			
Bestimmung von $\vartheta$ , Auswahl	1533	935,7	0,61
Bestimmung von $\hat{\varphi}$ , Auswahl	1390	1461,2	1,05
Bestimmung von $ \nabla ^{N^{(1)}}$ , Auswahl	1254	206,9	0,16
Generieren der Linien und der zugehörigen Gütekriterien	761	4212,4	5,53
Entfernen benachbarter ähnlicher Linien	258	216,7	0,84
Entfernen von Linien, die zu zu kleinen Segmenten führen würden	31	1,5	0,05
Schrittweises Realisieren der Trennlinien	27	279,8	10,30
	$\Sigma$	8052,3	

**Tabelle 5.2:** Laufzeitangaben für das Unterteilen eines Gebäudesegment. Jeder Arbeitsschritt wird für die jeweilige „Einheit“ einmal ausgeführt. Diese Einheiten sind auf Skelettpixelebene die noch nicht ausgeschlossenen Skelettpixel, auf Trennlinienebene die noch nicht ausgeschlossenen Trennlinienkandidaten.

In Tabelle 5.2 werden die in jedem Arbeitsschritt abzuwickelnden Einheiten, also entweder die Basispunkte in Form von Skelettpixeln oder die Trennlinienkandidaten, angegeben. Neben der Gesamtzeit für einen Arbeitsschritt lässt sich so auch die mittlere Laufzeit für eine Einheit ermitteln. Dadurch lässt sich die Laufzeit bei anders dimensionierten Ausgangsdaten abschätzen.

Die Laufzeiten unterstreichen den Sinn der schrittweisen Selektion der Trennlinienkandidaten. Insbesondere bei den aufwändigeren Schritten wie der Generierung der Linien (inkl. Bestimmung der Gütekriterien) oder der endgültigen Realisierung wirkt die Vorauswahl sehr beschleunigend.

Eine Gesamtlaufzeit von unter 10 s für ein großes Gebäudesegment ist auch bei ausschließlicher Nutzung zur automatisierten Vorselektion durchaus vertretbar. Das manuelle Einpassen der Gebäudetrennungen ist auch für geübte Benutzer kaum unter 2 min zu bewerkstelligen. Somit ist also auch bei einer automatisierten Vorselektion mit anschließender manueller Nachbearbeitung ein Zeitvorteil zu erwarten.





## 6 Nachbearbeitung der Gebäudekonturen

Das Oberflächenmodell stellt die Grundlage zur Ermittlung der Gebäudekonturen, wie sie in den vorangegangenen Kapiteln geschildert wurde, dar. Spektrale Information kommt bei der Unterteilung der Gebäude- und Vegetationssegmente wie auch bei der Aufteilung der Gebäudesegmente in kleinere Einheiten zum Tragen. Die Abgrenzung der Gebäude zur Umgebung (also keine Vegetation oder andere Gebäude) ist jedoch ausschließlich durch die Morphologie des DOM bestimmt. Wie in den Abschnitten 2.3.4 und 2.5.2 beschrieben, kann dem Oberflächenmodell gerade an solchen Grenzen durch das Stereoverfahren oft kein Höhenwert zugeordnet werden. Die interpolierten Höhenwerte sind nur eine grobe Schätzung der Realität. Die in Abschnitt 3.5 beschriebene Konturierung des DOM führt in solchen Fällen zu ungeraden oder „auslaufenden“ Konturlinien. Die Kanten der Gebäude im Orthobild hingegen sind von relativ guter Qualität. Gebäudekonturen erscheinen geradlinig und meist kontrastreich. Die Idee ist es nun, die durch das Oberflächenmodell bestimmten Gebäudekonturen mit Hilfe der Bildkanten eines Orthobildes zu verbessern. Das in diesem Kapitel als „Nachbearbeitung der Gebäudekonturen“ bezeichnete Vorgehen stellt eine Verschneidung der bisher ermittelten Konturen mit Bildkanten des Orthobildes dar.

Die im Rahmen der Konturüberarbeitung behandelten Teilschritte werden zunächst abschnittsweise vorgestellt (Abschnitt 6.2–6.5); in Abschnitt 6.6 werden die Einzelschritte dann im Kontext des Gesamtablaufs betrachtet. Die Ausgangskonturen der Gebäude sowie der aus einem Spektralkanal ermittelte Bildgradient stellen die Grundlage aller in diesem Kapitel vorgestellten Methoden dar. Sie werden in einem in Abschnitt 6.2 beschriebenen Vorverarbeitungsschritt für die nachfolgenden Schritte aufbereitet. In Abschnitt 6.3 werden zwei Optimierungsansätze zur iterativen Adaption der Gebäudekontur an Bildkanten präsentiert. Um bei der Adaption gebäudetypische Geometrien zu favorisieren, werden in Abschnitt 6.4 entsprechende Nebenbedingungen formuliert. Durch eine „dynamische Verwaltung“ der Stützpunkte des Umrisspolygons wird die optimale Anzahl zur Beschreibung der Hauskontur ermittelt. Auf diesen Aspekt wird in Abschnitt 6.5 eingegangen.

Nach der Vorstellung des Gesamtablaufs in Abschnitt 6.6 folgt eine Bewertung der Methoden in Bezug auf die damit erzielte Genauigkeit sowie die nötige Rechenzeit in Abschnitt 6.7. Auch die Reihenfolge von Konturnachbearbeitung und der im vorhergehenden Kapitel behandelten Unterteilung der Gebäudesegmente wird diskutiert.

## 6.1 Einführung und Literaturüberblick

Ansätze zur Gebäudeextraktion mit einer *vorhergehenden Liniendetektion* (s. auch Diskussion in Abschnitt 1.3 und 5.1) führen implizit zu geradlinigen Hauskonturen entlang entsprechender Bildkanten. Durch die Liniendetektion wird der Großteil der Gebäudeumrisse erfasst und liegt dann — zu meist unterteilt in kleinere Liniensegmente — zusammen mit gebäudefremden Linienstrukturen für eine nachfolgende Verarbeitung vor. Deren Ziel ist die Unterdrückung der gebäudefremden Kanten sowie die Zuordnung aller zu einer Gebäudekontur gehörigen Liniensegmente. Bei diesem schrittweisen Vorgehen *summieren sich Ungenauigkeiten und Fehler auf*: Beispielsweise tauchen Ungenauigkeiten in der Liniendetektion nicht nur im Endresultat auf, sondern können auch zu falschen Entscheidungen bei den nachfolgenden Schritten führen.

Als Vertreter der linienbasierten Vorgehensweise seien die bereits in Abschnitt 1.3.1 erwähnten Ansätze von KRISHNAMACHARI UND CHELLAPA (1996), SHUFELT (1997), YOON ET AL. (1999), SOHN UND DOWMAN (2001) genannt.

Mit den ebenfalls einleitend aufgeführten *modellbasierten Ansätzen* wird gebäudespezifisches Wissen in den Extraktionsprozess einbezogen. Je einfacher man das Gebäudemodell gestaltet, desto weniger Details können erfasst werden. Dafür ist die Schätzung der Modellparameter umso robuster gegenüber verrauschten Bilddaten. Diese mit der Komplexität parametrisierter Modelle einhergehende Koppelung von Vor- und Nachteilen entgegnet STEINHAGE (1997) bzw. FISCHER ET AL. (1998) mit einem generischen Modellierungsansatz. Mit einer Modell-Hierarchie von der Merkmalsebene bis zur Gebäudeebene erfolgt die Gebäudeextraktion daten- und modellgetrieben. Dabei geschieht die Hypothesenbildung und -verifikation gekoppelt sowohl im zweidimensionalen Bildraum als auch im dreidimensionalen Modellraum.

Für das in Abschnitt 1.4.1 definierte Ziel, möglichst wenige Annahmen bezüglich der Gebäudegeometrie zu treffen, eignen sich modellbasierte Ansätze nur bei sehr allgemeinen Gebäudemodellen oder einer sehr großen Modellbibliothek. Auf die damit verbundenen Nachteile, insbesondere der fehleranfälligeren Parameterschätzung, wurde bereits hingewiesen (s. Abschnitt 1.3.1).

Unter der Nebenbedingung, rechte Winkel einzuführen, vereinfachen WEIDNER UND FÖRSTNER (1995) ein aus dem DOM gewonnenes Gebäude-Konturpolygon durch Nutzung der minimalen Beschreibungslänge (MDL) des Polygons. Die Beschreibungslänge ergibt sich aus der Übereinstimmung des aktuellen Konturpolygons mit Prismenmodellen unterschiedlicher Komplexität und der Komplexität des jeweiligen Modells. Die so gewonnenen Konturen werden ein weiteres Mal überarbeitet, um geometrische Beziehungen wie auch die Übereinstimmung zu den ursprünglichen Konturen der Gebäudesegmente zu gewährleisten. Da in diesem Ansatz kein festes Gebäudemodell vorgegeben ist, kann nicht von einer Einschränkung bezüglich der Gebäudeform gesprochen werden. Datengrundlage für den Ansatz von WEIDNER UND FÖRSTNER (1995) ist ausschließlich ein DOM. Dadurch kann das Konturpolygon nicht anhand weiterer Bilddaten modifiziert werden. Die Überarbeitung des Konturpolygons erfolgt nur anhand geometrischer Kriterien ohne Zuhilfenahme

von Bildinformation. Auch die Übereinstimmung des modifizierten Konturpolygons mit dem DOM wird nicht mehr überprüft.

Eine Modifikation von Konturlinien *unter Beachtung von Bildinformation* wird mit dem in der Bildverarbeitung als „Aktive Kontur“ bezeichneten Konzept ermöglicht (s. Abschnitt 6.3). Liegt bereits ein initialer Gebäudeumriss vor, kann eine aufwändige Zuordnung der Bildkanten zu einem Gebäudeobjekt entfallen (linienbasierte Verfahren). Eine *Angleichung der Konturlinie an Bildkanten* in der Nachbarschaft bietet sich in einem solchen Fall an. In ihrer ursprünglichen Implementierung sind Aktive Konturen allerdings schlecht für die Erfassung der meist rechtwinklig angeordneten Hausumrisse geeignet. Um solche Diskontinuitäten der Konturlinienkrümmung modellieren zu können, müssen zusätzliche Vorkehrungen getroffen werden.

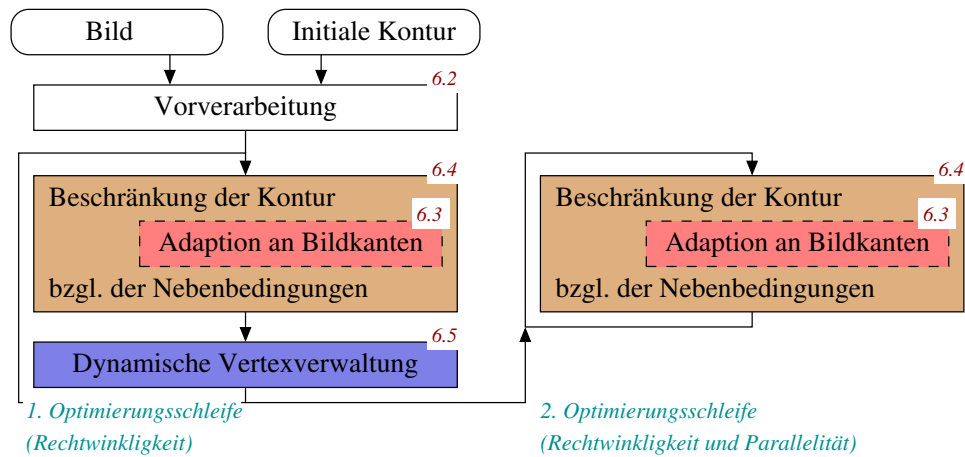
RÜTHER ET AL. (2002) passen Gebäudeumrisse mittels Aktiver Konturen an nahe Bildkanten an. Die Konstellation rechter Winkel soll bei dem Optimierungsprozess durch *Erweitern der Energiefunktion* um einen entsprechenden Geometrie-Term favorisiert werden. Da von einer nahezu homogenen Pixelverteilung innerhalb der Dachfläche ausgegangen wird, ist der Ansatz auf *einfache freistehende Häuser* beschränkt.

YOON ET AL. (1999) nutzen ein Aktives Konturmodell zur Selektion von bereits detektierten Liniensegmenten. Die Aktive Kontur wird durch Hinzufügen eines Linien-Terms so modifiziert, dass eine Adaption an rechtwinklige Strukturen möglich wird. Liniensegmente entlang der so detektierten Konturlinie werden nun, wie bei linienbasierten Ansätzen üblich (s. vorhergehende Anmerkungen diesbezüglich), verlängert bzw. verkürzt oder es werden neue Linien generiert. Durch einen Vergleich aller möglichen Konturhypothesen wird eine endgültige Auswahl getroffen.

In dem bereits einleitend in Kapitel 1.3.2 vorgestellten Ansatz von VESTRI UND DEVERNAY (2001) werden die Eckpunkte bereits extrahierter planarer Flächen eines DOM ermittelt. Aus einer Aneinanderreihung der Eckpunkte resultiert dann ein Gebäudepolyeder. Bei dem Aneinanderreihungsprozess wird bereits die Position der Polygonstützpunkte in Bezug auf die Nachbarn und den Verlauf der Konturlinie optimiert. Die Verteilung der Polygonvertices wird also *vor und unabhängig von der darauf folgenden Feinkorrektur* der Konturpolygone hinsichtlich ihrer Geometrie behandelt. Die Feinkorrektur soll die Rechtwinkligkeit oder Geradlinigkeit der Polygonsegmente wie auch deren Abstand zum Ausgangspolygon optimieren. Die Optimierung geschieht durch einen Minimierungsprozess, in den die Nebenbedingungen additiv einfließen. Die Nebenbedingungen sind entsprechend der in Abschnitt 6.4.1 folgenden Diskussion „weich“ formuliert. Die Optimierung betrifft *ausschließlich die Positionen der Vertices* im 3-D-Raum; Bild- bzw. DOM-Merkmale werden dabei nicht mehr in Betracht gezogen.

Der in diesem Kapitel vorgestellte Ansatz zur Überarbeitung der Hauskonturen ist schematisch in Abbildung 6.1 wiedergegeben. Der Ansatz orientiert sich an dem Konzept der Aktiven Konturen. Auf die eine Aktive Kontur auszeichnenden Energieterme zur Regularisierung wird jedoch verzichtet. Die Konturlinie wird als einfaches *Polygon* modelliert. Das Ausprägen rechter Winkel ist so ohne Weiteres möglich. Die Regularisierung erfolgt abgekoppelt vom Optimierungsprozess in Form einer

„dynamischen Verwaltung der Polygonvertices“. Dadurch wird zugleich die Anzahl der Vertices, die zur Beschreibung der Konturlinie benötigt wird, minimiert.



**Abbildung 6.1:** Vereinfachte Darstellung zur Nachbearbeitung der Gebäudekonturen.

Die bei modellbasierten Ansätzen übliche Beschränkung bezüglich der Gebäudekomplexität (bzw. die zu grobe Extraktion bei komplexeren Häusern) ist durch die polygonale Erfassung der Gebäudekontur nicht nötig. Eine Reduktion des Einsatzgebietes auf quaderförmige Fabriken oder bestimmte Dachtypen wird somit vermieden. Prinzipiell sind beliebig komplexe Umrisse erfassbar. Gebäudetypische Geometrien werden allerdings durch entsprechend formulierte *Nebenbedingungen* bevorzugt. Die Umsetzung der Nebenbedingungen ist in den Optimierungsprozess integriert und erfolgt gekoppelt an den in jeder Iteration vollzogenen Adaptionsschritt. In Abbildung 6.1 ist diese Koppelung von geometrischen Nebenbedingungen und Adaption durch ineinander geschachtelte Boxen dargestellt.

Wie Abbildung 6.1 zu entnehmen ist, ist die Nachbearbeitung der Gebäudekonturen aufgeteilt in *zwei Schleifen*. In der ersten Schleife erfolgt eine Adaption an Bildkanten unter der Nebenbedingung, an geeigneten Vertices rechte Winkel einzuführen. Zudem wird die Zahl der Polygonvertices minimiert. Die zweite Schleife gilt als Feinkorrektur, in der zusätzliche Nebenbedingungen eine gleiche Ausrichtung geeigneter Polygonsegmente bewirken.

Bei Aktiven Konturen erfolgt die Anpassung an umgebende Bildkanten üblicherweise auf Grundlage der Gradientenstärke. Bei dem in diesem Kapitel vorgestellten Ansatz wird zusätzlich auf die *Gradientenorientierung* zurückgegriffen. Polygonsegmente, bei denen die der Linie zugeordneten Pixel eine einheitliche Orientierung aufweisen, erhalten eine höhere Bewertung. Damit wird die Erfassung linearer Strukturen, wie sie für Hausumrisse typisch sind, gefördert.

Durch den schrittweise ablaufenden Adaptionsprozess wird die mit linienbasierten Ansätzen einhergehende Fehlerkumulation weitestgehend umgangen. Kleinere Abweichungen der Ausgangskontur, hervorgerufen durch Ungenauigkeiten oder Fehler der Vorgängerschritte (Kapitel 3 und 4) oder gar des Ausgangsprodukts DOM, können so kompensiert werden. Diese Robustheit des adaptiven

Vorgehens geht zu Lasten der Laufzeit, da in jeder Iteration Zugriffe auf den Bildgradienten nötig sind. Die Abstraktion linienbasierter Vorgehensweisen hingegen ermöglicht ein rein koordinatenbasiertes Arbeiten. Sind die relevanten Umrisslinien der Gebäude einmal extrahiert, so muss nicht mehr auf die Bildinhalte zurückgegriffen werden. Die geometrische Zuordnung der Linien zu einzelner Gebäuden kann anhand der (möglicherweise mit Bildinformation attribuierten) Linienkoordinaten erfolgen.

Der Einfachheit halber wird bei der Verbesserung der Gebäudekontur auf die Integration von Kontextwissen, wie z. B. die Ausrichtung bereits extrahierter Häuser oder des umgebenden Gebäudekomplexes, verzichtet. Die Gebäudeobjekte wie auch die Umrisse von deren Innenhöfen werden unabhängig voneinander behandelt.

## 6.2 Vorverarbeitung der Bilddaten

Grundlage der Verbesserung einer Hauskontur ist zum einen eine initiale Konturlinie. Wie bereits erwähnt, geht sie in Form eines geschlossenen Polygons bereits aus den Arbeitsschritten der vorhergehenden Kapitel hervor. Zum anderen dient der Bildgradient als Basis für eine Adaption des Gebäudepolygons. Aus diesen Ausgangsdaten werden im Folgenden weitere Basisdaten generiert.

### 6.2.1 Polygonvereinfachung

Letztendlich stammt das Ausgangspolygon aus der in Abschnitt 3.5 beschriebenen Konturierung eines normierten Höhenmodells. Der geschlossene Polygonzug, der die Gebäudekontur repräsentiert, wird im Folgenden mit  $V$  bezeichnet

$$V = \{v_i, i = 1, \dots, n\} = \{(x_i, y_i)\}. \quad (6.1)$$

Das Polygon habe  $n$  Stützstellen oder Vertices.

Da die Konturierung pixelgenau erfolgt, ist dies mit einer hohen Anzahl an Stützstellen verbunden. Die Komplexität der in Abschnitt 6.3 vorgestellten Adaption des Umrisspolygons an Bildkanten bzw. der nachfolgend eingeführten Nebenbedingungen steigt jedoch mit der Anzahl der Stützstellen. Zudem ist das Detailreichtum in Relation zur Genauigkeit der initialen Konturlinie deutlich zu hoch. Aus diesem Grund wird das Polygon vereinfacht, d. h. die Zahl der Stützstellen wird auf ein Minimum reduziert ohne dabei zu sehr die Gestalt des Polygons zu verändern.

Grundlage dafür ist der Algorithmus von DOUGLAS UND PEUCKER (1973). In jedem Schritt wird versucht, eine Reihe aufeinanderfolgender Polygonlinien durch ein Liniensegment zu approximieren. Zur Beurteilung wird der Punkt der Liniensequenz gesucht, der am weitesten von der einzelnen Linie entfernt ist. Ist die Entfernung kleiner als ein vorzugebender Schwellwert (Approximationsfehler), so wird die Linienapproximation akzeptiert. Andernfalls wird der Algorithmus rekursiv auf die zwei

Linienzüge von und bis zu dem gewählten Punkt angewendet.

Der so genannte Douglas-Peucker-Algorithmus ist heuristisch und nicht optimal bezüglich der Approximation. Verglichen mit anderen heuristischen Vereinfachungsansätzen führt der Douglas-Peucker-Algorithmus jedoch zu den subjektiv wie objektiv besten Approximationsergebnissen (MC-MASTER 1987).

Der Douglas-Peucker-Algorithmus ist auf Polygonsegmenten definiert. Zur Anwendung auf geschlossene Polygone, wie sie in dieser Arbeit vorkommen, wird ein beliebiger Ausgangsvertex gewählt. Durch den Vertex mit der maximalen Entfernung dazu sind zwei Polygonsegmente definiert. Diese werden dann nach dem beschriebenen Vorgehen abgearbeitet. In Abbildung 6.2(a) ist einer Bildszene eine initiale Gebäudekontur überlagert; Abbildung 6.2(b) zeigt das vereinfachte Konturpolygon. Gegenüber dem Ausgangspolygon mit 318 Stützstellen kommt das vereinfachte Polygon mit 20 Vertices aus.

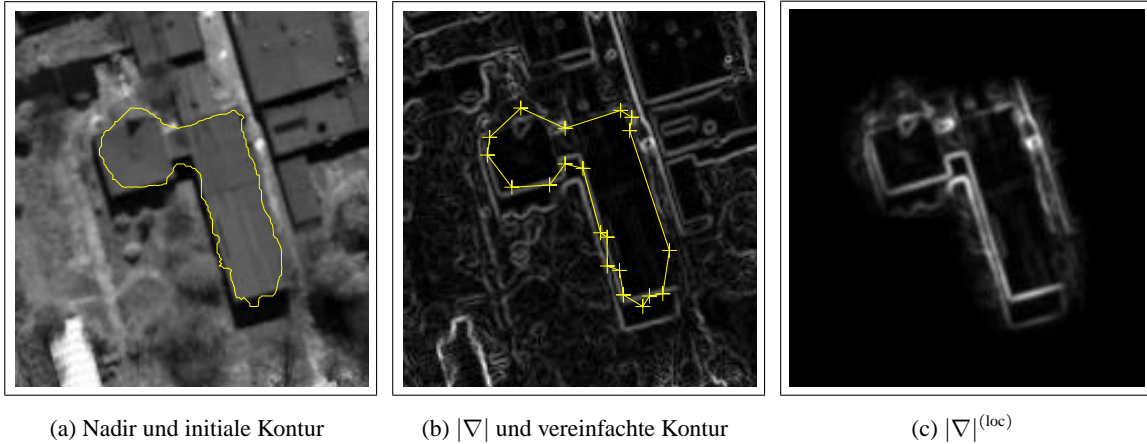
Die Polygonvereinfachung wird auch zu der in Abschnitt 6.5 beschriebenen Regularisierung des Umrisspolygons herangezogen. Dabei gibt es allerdings Stützpunkte, die bei einer Vereinfachung nicht entfallen dürfen. Deshalb wird der Douglas-Peucker-Algorithmus so modifiziert, dass die Vereinfachung auf den spezifizierten Punkten aufbaut. Eine Vereinfachung findet dann nur noch auf den Segmenten zwischen zwei fixen Stützstellen statt.

### 6.2.2 Lokale Gradientenstärke

Die Konturlinien der Gebäude zeichnen sich im Bild meist durch starke Grauwertwechsel ab. Wie schon im vorhergehenden Kapitel stellt sich deswegen der Bildgradient als zentrales Merkmal für Gebäudekanten dar. Entlang einer geradlinigen Gebäudekante liegt eine hohe Gradientenstärke bei relativ konstanter Gradientenorientierung vor. Der Bildgradient berechnet sich nach Gleichung 1.1 in Abschnitt 1.5.3. Grundlage dafür ist einer der Bildkanäle, vorzugsweise der meist höherauflösende Nadir.

Die Ausgangspolygone der Gebäude werden durch die in diesem Kapitel beschriebenen Methoden so modifiziert, dass die Gradientenstärke entlang der Polygonlinie maximal wird. Das initiale Gebäudepolygon ist nicht frei gewählt, sondern bereits eine Abschätzung des realen Umrisses. Durch dieses Vorwissen können bereits weite Bereiche des Bildes als mögliche Polygonpositionen ausgeschlossen werden. Neben einer beschleunigten Konvergenz der Polygonadaptation hat dies vor allem eine höhere Robustheit gegenüber objektfremden Bildkanten zur Folge. Idealerweise sollte eine so überarbeitete Gradientenstärke nur die Kanten des aktuellen Gebäudes erfassen. Implementiert ist dies durch das Unterdrücken der Gradientenstärke für Bildpunkte fern von dem anfänglichen Gebäudepolygon. Resultat ist eine „lokale Gradientenstärke“  $|\nabla|^{(loc)}$ , die sich im Einzelnen wie folgt berechnet:

Aus der initialen Gebäudekontur werden die zugehörigen Pixelpositionen des gesamten Polygonzugs ermittelt. Daraus resultiert ein (binäres) Konturbild  $F_V$ , das mit einem breiten Gauß'schen Kern  $G^{(0)}$  gefaltet wird, was einer Glättung entspricht. Als weiteres Teilprodukt wird die Gradientenstär-



**Abbildung 6.2:** Bildbeispiel zur Polygonvereinfachung und Ermittlung der lokalen Gradientenstärke  $|\nabla|^{(loc)}$ . Das Ausgangsbild ist überlagert mit der initialen Gebäudekontur (a). Die zugehörige Gradientenstärke überlagert mit dem vereinfachten Gebäudepolygon (b) sowie die daraus abgeleitete lokale Gradientenstärke (c).

ke herangezogen. Sie wird auf einen maximalen Wert  $|\nabla|_{\max}$  begrenzt und mit einem schmalen Gauß'schen Kern  $G^{(1)}$  gefaltet. Die lokale Gradientenstärke ergibt sich dann durch Multiplikation der beiden Faltungsergebnisse:

$$|\nabla|^{(loc)} = \left( F_V * G^{(0)} \right) \cdot \left( \min(|\nabla|, |\nabla|_{\max}) * G^{(1)} \right) \quad (6.2)$$

Die Multiplikation mit dem „verschmierten“ Konturbild bewirkt bei der Durchführung der Konturüberarbeitung die erwähnte Unterdrückung von Bildkanten fern des aktuell betrachteten Konturpolygons. Die Glättungen sind nötig, um die schrittweise durchgeführte Adaption überhaupt erst zu ermöglichen. Wie in Abschnitt 6.3 noch ausführlicher beschrieben wird, muss zur Adaption an eine Bildkante ein Anstieg der Gradientenstärke bereits in einer Umgebung der Kante vorhanden sein. Nur so kann sich das Umrisspolygon auf den Ort mit der höchsten Gradientenstärke zu bewegen. In Abbildung 6.2(c) ist exemplarisch die aus der Gradientenstärke und der initialen Kontur abgeleitete lokale Gradientenstärke als Grauwertdarstellung wiedergegeben.

### 6.3 Adaption an Bildkanten

Die Adaption der Konturpolygons an Bildkanten ist angelehnt an Aktive Konturmodelle, die durch KASS ET AL. (1988) auch unter dem Namen „Snakes“ bekannt wurden. Eine Aktive Kontur ist eine parametrische Kurve („Spline“), der eine Energie zugeschrieben wird. Diese Energie setzt sich zusammen aus internen Energien, die der Regularisierung der Kurve dienen, und der externen Energie, die die Kurve zum betrachteten Objekt führen soll. Die Anpassung an das Objekt er-

folgt über eine iterative Minimierung der Gesamtenergie bis ein Gleichgewichtszustand, d. h. ein lokales Energie-Minimum, vorliegt. KASS ET AL. (1988) stellen zur Minimierung der Energiefunktion Euler-Lagrange-Gleichungen auf, die numerisch mittels Variationsrechnung gelöst werden. Für die interne Energie werden meist die Dehnung und die Krümmung in den Stützpunkten verwendet. Bei fehlenden externen Kräften würden sich (in der Original-Implementierung geschlossener Snakes von KASS ET AL. (1988)) die Stützpunkte gleichmäßig in Form eines Kreises verteilen, der nach und nach in sich zusammenfällt. Als externe Energie kommt meist der negative Gradient des Bildes (abgegriffen an den jeweiligen Stützstellen) zum Einsatz. Bildpunkte mit hohem Gradienten, also Bildkanten, üben dann eine besonders große „Anziehungskraft“ auf die Kontur aus.

Die Implementierung als Spline-Kurve führt zu einer glatten und differenzierbaren Konturlinie. Bei der Modellierung von Gebäudeumrissen ist zu beachten, dass rechte Winkel ein essentielles Merkmal der meisten Häuser sind. Diese Unstetigkeiten in der Konturkrümmung widersprechen der Glattheitsannahme der internen Energie einer Aktiven Kontur. Sie können nur durch eine Ausnahmebehandlung der betreffenden Stützpunkte erfasst werden. So wird z. B. bei WILLIAMS UND SHAH (1992) für die als Eckpunkte erkannten Stützstellen das Krümmungskriterium (durch Multiplikation mit 0) aufgehoben. Die in dieser Arbeit eingeschlagene Strategie ist jedoch, auf die internen Energien gänzlich zu verzichten und die Kontur anderweitig zu regularisieren (s. Abschnitt 6.5). Wie in Gleichung 6.1 dargestellt, wird die Kontur als geschlossenes Polygon repräsentiert. Da Hauswände und Dachkanten im Normalfall geradlinig sind, ist die Modellierung der Konturkurve als Polygon ausreichend (vorausgesetzt, die Polygonstützstellen erfassen die Ecken der Gebäudekontur).

Im Folgenden wird die Adaption der Polygonstützpunkte an Bildkanten behandelt. Bildkanten sind in diesem Zusammenhang die Regionen mit hoher lokaler Gradientenstärke; die lokale Gradientenstärke wird vorab entsprechend dem letzten Abschnitt ermittelt. In Abschnitt 6.3.1 folgt die Definition eines Gütemaßes, das auf der lokalen Gradientenstärke aufbaut. In Abschnitt 6.3.2 werden zwei Ansätze zur Optimierung dieser Gütefunktion diskutiert.

Die Zahl der Stützpunkte wird innerhalb dieses Abschnitts als fix betrachtet, ebenso bleibt die Ausprägung einer gebäudespezifischen Geometrie hier unbeachtet. Auf diese Aspekte wird in den nachfolgenden Abschnitten eingegangen.

### 6.3.1 Ermittlung der Polygongüte

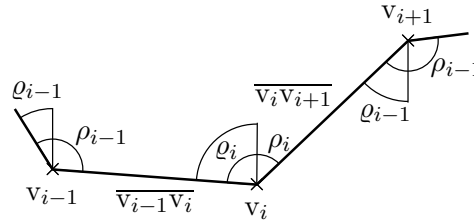
Zur Annäherung des Konturpolygons an gute Bildkanten muss ein Gütemaß zur Beurteilung der Polygongüte definiert werden. Dies geschieht analog zur der gewichtet gemittelten Gradientenstärke  $|\overline{\nabla}|$ , die im vorhergehenden Kapitel zur Beurteilung von Trennlinien eingeführt wurde (s. Gleichung 5.3). Das Polygonsegment zwischen den zwei Stützpunkten  $v_i$  und  $v_{i+1}$ ,  $\overline{v_i v_{i+1}}$ , wird entsprechend der lokalen Gradientenstärke bewertet. Genau wie bei der Mittelung der Gradientenstärke in Abschnitt 5.3 sollen auch hier die Gradienten von Pixeln, deren Orientierung nicht mit der des Polygonsegments übereinstimmt, unterdrückt werden. Dazu wird zuerst die Orientierung eines Po-



lygonsegments  $\overline{v_i v_{i+1}}$  definiert:

$$\varrho_i = \left( \arctan \left( \frac{x_{i+1} - x_i}{y_{i+1} - y_i} \right) + \pi \right) \bmod \pi \quad (6.3)$$

In Abbildung 6.3 sind die einem Polygonzug zugeordneten Orientierungen  $\varrho$  exemplarisch dargestellt.



**Abbildung 6.3:** Schematische Darstellung zur Nomenklatur der geometrischen Beziehungen innerhalb eines Polygons. Mit  $\rho$  sei der von zwei Polygonsegmenten eingeschlossene Winkel (also jeweils  $\leq \pi$ ), mit  $\varrho$  die absolute Orientierung eines Polygonsegments (ebenfalls in  $[0, \pi]$ ) bezeichnet.

Die Menge der Bildpunkte, die dem Polygonsegment  $\overline{v_i v_{i+1}}$  zugeordnet sind, sei mit  $\mathcal{V}_i$  bezeichnet. Für jeden dieser Punkte  $j \in \mathcal{V}_i$  ergibt sich nun, analog zu Gleichung 5.2, der Unterschied von Gradientenorientierung  $\varphi$  und Richtung des Polygonsegments als

$$\Delta\varphi_j^{(\mathcal{V}_i)} = \min(|\varrho_i - ((\varphi_j + 3/2\pi) \bmod \pi)|, \pi - |\varrho_i - ((\varphi_j + 3/2\pi) \bmod \pi)|) . \quad (6.4)$$

Anhand dieser Winkeldifferenz wird die Gradientenstärke der betroffenen Pixel gewichtet und führt so zu einer gemittelten Gradientenstärke für ein Polygonsegment; sie stellt das Gütemaß  $q_{i,i+1}^{(V)}$  zur Beurteilung eines Polygonsegments dar.

$$q_{i,i+1}^{(V)} = \overline{|\nabla|}^{(v_i)} = \frac{1}{|\mathcal{V}_i|} \sum_{j \in \mathcal{V}_i} |\nabla|_j^{(\text{loc})} \left( 1 - \frac{\min(\Delta\varphi_j^{(\mathcal{V}_i)}, \Delta\varphi_{\max})}{\Delta\varphi_{\max}} \right) \quad (6.5)$$

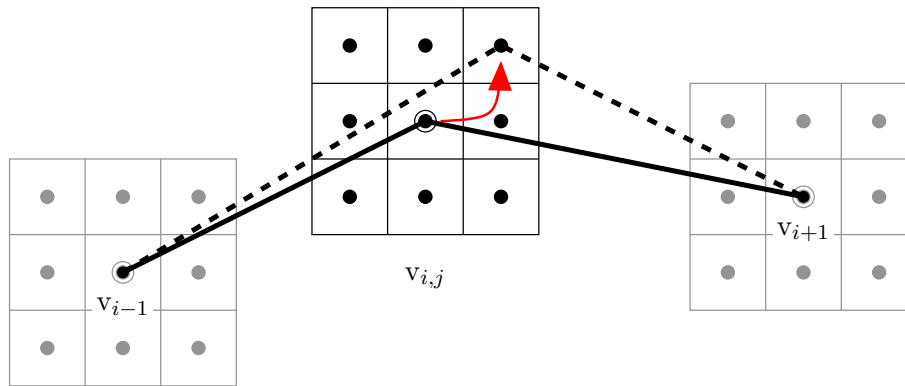
Mittels des Schwellwerts  $\Delta\varphi_{\max}$  wird festgelegt, wie weit  $\varphi$  und  $\varrho$  maximal differieren dürfen, um noch in die Wertung einzufließen.

Die Gradientenstärke für ein gesamtes Polygon ergibt sich dann aus der Summe der Gütewerte der Polygonsegmente:

$$Q^{(V)} = \overline{|\nabla|}^{(V)} = \sum_{i=1}^n \overline{|\nabla|}^{(v_i)} \quad (6.6)$$

### 6.3.2 Optimierung der Konturgüte

Im Folgenden wird vorausgesetzt, dass zur Optimierung des Konturpolygons die diskreten Koordinaten des Bildrasters herangezogen werden. Da die Adaption iterativ erfolgt, genügt es, die Beweglichkeit der Stützpunkte auf die unmittelbare Nachbarschaft zu begrenzen (z. B. ein  $3 \times 3$  Pixel großes Fenster wie in Abbildung 6.4 dargestellt). Diese Begrenzung des Suchraums macht eine (numerische) Optimierung der Kontur überhaupt erst praktikabel. Doch selbst bei  $n$  Polygonstützstellen und einem Suchfenster von  $m$  Pixeln benötigt eine erschöpfende Suche  $m^n$  Vergleiche, wobei jeder Vergleich die Ermittlung der Konturgüte voraussetzt. Schon bei 20 Stützstellen und einem  $5 \times 5$  Suchfenster wären  $25^{20} \approx 10^{28}$  Vergleiche notwendig. Daher ist eine weitere Einschränkung des Suchraums für einen praktischen Einsatz unabdingbar. Zwei Herangehensweisen hierzu werden im Folgenden diskutiert. Die Betrachtung der Nachbarschaft eines Vertex  $v_i$  wird durch einen zusätzlichen Index  $j = 1, \dots, m$  kenntlich gemacht ( $v_{i,j}$ ).



**Abbildung 6.4:** Schematische Darstellung zur Adaption der Polygonstützpunkte auf dem diskreten Bildraster. In einem Suchfenster um Stützpunkt  $v_i$  (die Nachbarnpixel werden mit  $v_{i,j}$  bezeichnet) wird derjenige Bildpunkt zum neuen Stützpunkt, der zur besten Polygonbewertung führt.

### Globale Optimierung mittels Dynamischer Programmierung

Wie bereits erwähnt, sollen Polygonstützpunkte so verschoben werden, dass kumuliert über die gesamte Polygonlänge die höchste Güte erreicht wird. Im Gegensatz zur erschöpfenden Suche, bei der alle (bezogen auf das Suchfenster um jeden Vertex) Konstellationen der Stützpunkte betrachtet werden, wird bei der Dynamischen Programmierung die Auswahl eingeschränkt (BELLMAN 1957). Die optimale Vertex-Konstellation wird dabei schrittweise für jeden Stützpunkt ermittelt. Die Suche nach dem optimalen Pfad vom Anfangsvertex zum aktuellen Vertex kann so auf die Verbindung zu den Vorgängern reduziert werden. Die Adaption Aktiver Konturen mittels Dynamischer Programmierung wurde erstmals durch AMINI ET AL. (1988) und YUILLE ET AL. (1989) vorgestellt.

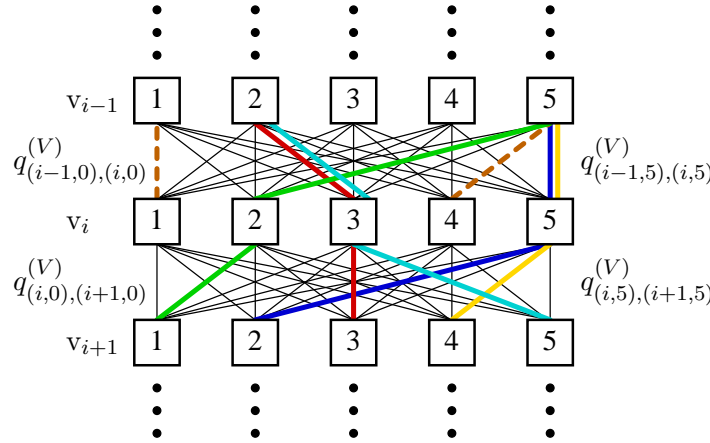
Für den Fall von  $n$  Polygonstützstellen und einer Nachbarschaft von  $m = 5$  Rasterpunkten ergibt sich die in Abbildung 6.5 schematisch dargestellte Situation. Den Stützpunktkandidaten wird

eine Güte  $Q^{(V)}$  zugeschrieben, die die kumulierte Kantengüte vom Startvertex zum aktuellen Vertex widerspiegelt. Um die Stützpunktkandidaten  $v_{i,j}$ ,  $i = 1, \dots, n$ ,  $j = 1, \dots, m$  zu bewerten, werden zur bereits berechneten Bewertung der Vorgänger jeweils die Güte der betreffenden Verbindung  $q_{(i-1,k),(i,j)}^{(V)}$  addiert. Dann wird für alle Nachbarn eines Stützpunkts der Vorgänger mit der besten Bewertung,  $v_{i-1,\tilde{k}}$ , gespeichert. Dieser ergibt sich aus

$$\tilde{k} = \operatorname{argmax}_k q_{(i-1,k),(i,j)}^{(V)} + Q_{i-1,k}^{(V)}, \quad (6.7)$$

wobei  $Q_{i-1,k}^{(V)}$  die kumulierte Bewertung bis zum Vertexkandidaten  $v_{i-1,k}$  darstellt. Die Bewertung des Vertexkandidaten  $v_{i,j}$  ergibt sich somit als

$$Q_{i,j}^{(V)} = q_{(i-1,\tilde{k}),(i,j)}^{(V)} + Q_{i-1,\tilde{k}}^{(V)}. \quad (6.8)$$



**Abbildung 6.5:** Schematische Darstellung zur Optimierung der Polygonstützpunkte nach dem Prinzip der Dynamischen Programmierung. Zeilenweise sind für jeden Vertex die dem Suchfenster entsprechenden Nachbarn aufgetragen. Die Verbindungen zwischen den Stützpunkten kennzeichnen das resultierende Polygonsegment. Die farbig dargestellten Verbindungen geben exemplarisch den jeweils besten Pfad zu den Vorgängervertices wieder.

Gegenüber der erschöpfenden Suche wird also für jeden Vertex nur *ein* Pfad zum bestbewerteten Vorgänger gespeichert. In Abbildung 6.5 sind diese Pfade farbig eingezeichnet.  $v_{i-1,5}$  stellt in diesem Beispiel den bestbewerteten Vorgänger des Vertexkandidaten  $v_{i,5}$  dar. Alle nachfolgenden Vertices, für die  $v_{i,5}$  der bestbewertete Vorgänger ist ( $v_{i+1,2}$  und  $v_{i+1,4}$ ), haben somit gleichzeitig  $v_{i-1,5}$  als (Vor-)Vorgänger. Braun gestrichelt eingezeichnet sind zwei Verbindungen ( $\overline{v_{i-1,1}v_{i,1}}$  und  $\overline{v_{i-1,5}v_{i,4}}$ ), die zwar jeweils eine bestbewertete Verbindung darstellen, doch bereits im nächsten Bewertungsschritt entfallen. Grund dafür ist, dass  $v_{i,1}$  und  $v_{i,4}$  für keinen der Folgekandidaten als bestbewerteter Vorgänger gelten.

Ist die Bewertung der letzten Stützpunktkandidaten  $v_{n,m}$  abgeschlossen, so resultiert aus dem am

besten bewerteten Kandidaten die Polygondeformation, die zur höchsten kumulierten Kantengüte führt. Zu dem besten Endvertex ist nämlich der beste Vorgänger kandidat gespeichert, zu diesem wiederum der beste Vorgänger, usw. bis der Startvertex erreicht wird. Die Modellierung eines geschlossenen Polygons kann dadurch erreicht werden, dass der Startstützpunkt als zusätzlicher Vertex nach dem letzten Stützpunkt eingefügt wird.

Im Vergleich zur erschöpfenden Suche wird der Aufwand an Vergleichsoperationen von  $O(m^n)$  auf  $O(nm^2)$  reduziert. Die Suche der optimalen Konstellation mittels dynamischer Programmierung beruht auf dem Optimalitätsprinzip. Danach hat eine optimale Strategie die Eigenschaft, dass unabhängig vom Anfangszustand und den Anfangsentscheidungen die folgenden Entscheidungen wieder eine optimale Strategie bilden. Damit dies gilt, müssen für die Gütefunktionen bestimmte Monotoniebedingungen erfüllt sein (BELLMAN UND KALABA 1965). Für die verwendete additive Kantengütefunktion ist dies der Fall.

### Lokale Optimierung

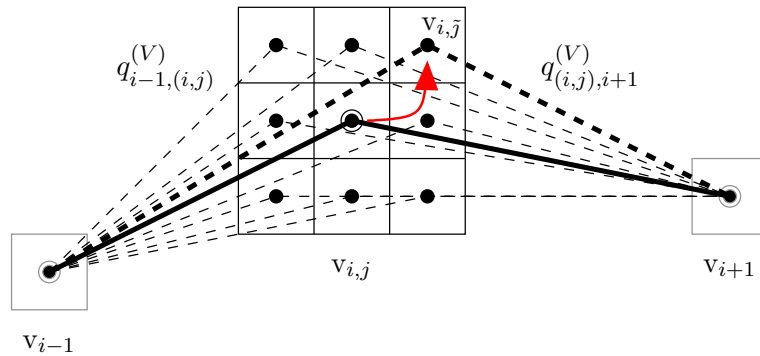
Bei einer lokalen Optimierung des Konturpolygons (ein so genannter „Greedy“-Algorithmus, s. z. B. CORMEN ET AL. (1990, Kapitel 17)) wird der Anspruch, eine global optimale Lösung zu finden, zugunsten einer erheblichen Einsparung an Rechenaufwand aufgegeben. Die Polygonstützpunkte werden unabhängig voneinander so bewegt, dass lokal für jeden Vertex die Kantengüte maximal wird. Anhand dieser lokal optimalen Wahl wird eine Annäherung an eine global optimale Lösung angestrebt. Die ist zwar keineswegs garantiert, doch sind die lokal getroffenen Entscheidungen meist auch global sehr gute Lösungen.

Ebenso wie bei der Dynamischen Programmierung wird die Optimierung des Polygons schrittweise für jeden Stützpunkt durchgeführt. Dabei werden die Positionen des Vorgängers und des Nachfolgers als fix angenommen. Nun wird derjenige Bildpunkt im Suchfenster als neuer Stützpunkt gewählt, der die Kantengüte maximiert. Da ein Polygonstützpunkt immer zwei weitere Vertices verbindet, werden diese beiden Verbindungen in die Berechnung der Kantengüte einbezogen. Für die Stützpunktkandidaten  $v_{i,j}$  wird die Kantengüte der Segmente  $\overline{v_{i-1}v_{i,j}}$  und  $\overline{v_{i,j}v_{i+1}}$  addiert (s. auch Abbildung 6.6):

$$\tilde{j} = \operatorname{argmax}_j q_{i-1,(i,j)}^{(V)} + q_{(i,j),i+1}^{(V)} \quad (6.9)$$

Die Entscheidung fällt also auf den Stützpunktkandidaten  $v_{i,\tilde{j}}$  mit der größten Kantengüte zum Vorgänger und Nachfolger.

Durch die Beschränkung auf die lokal beste Wahl kommt dieses Vorgehen mit einem  $O(nm)$  Rechenaufwand aus (die Vergleichsoperationen betreffend). Aufbauend auf einer lokalen Optimierung implementieren WILLIAMS UND SHAH (1992) Aktive Konturen. Für die vorgestellte Polygonadaptation sind sowohl ein lokaler wie auch ein globaler Optimierungsansatz mittels Dynamischer Programmierung realisiert. Eine Gegenüberstellung in Bezug auf die damit erreichbaren Genauig-



**Abbildung 6.6:** Schematische Darstellung zur lokalen Optimierung der Polygonstützpunkte. Jeder der Vertexkandidaten  $v_{i,j}$  wird anhand der Güte der Verbindung zum Vorgänger  $v_{i-1}$ ,  $q_{i-1,(i,j)}^{(V)}$ , und zum Nachfolger  $v_{i+1}$ ,  $q_{(i,j),i+1}^{(V)}$ , bewertet. Vorgänger und Nachfolger werden dabei als fix betrachtet.

keiten und die benötigten Laufzeiten folgt in Abschnitt 6.7.

## 6.4 Berücksichtigung gebäudetypischer Geometrien

Die im letzten Abschnitt beschriebene Adaption an Bildkanten verläuft rein datengetrieben. Nur die durch den Bildgradienten gegebene Güte der Polygonsegmente ist dafür ausschlaggebend. Jedoch zeichnet sich üblicherweise nicht die gesamte Hauskontur durch eine hohe Gradientenstärke aus. Beispielsweise kann der Kontrast von Bildkanten durch Schatten oder durch die Interpolation im Rahmen der Orthokorrektur (s. Abschnitt 2.3.5 und 2.5.1) stark reduziert werden. Außerdem wird im Fall einer initialen Kontur, die weit vom realen Hausumriss entfernt ist,  $|\nabla|^{(loc)}$  an den eigentlichen Hauskanten abgeschwächt (s. z. B. dritter Vertex von links in Abb. 6.2(b)).

Die Integration von externem Wissen kann zu einer größeren Robustheit gegenüber solchen Einflüssen führen. Dadurch können gewisse Polygonkonstellationen favorisiert bzw. andere vermieden werden. Da ausschließlich die Konturen von Gebäuden behandelt werden, kommen als externes Wissen Annahmen über möglichst typische Ausprägungen der Gebäudekontur in Betracht. So weist ein Großteil der Häuser rechte Winkel auf; auch sind gegenüberliegende Hauswände meist parallel ausgerichtet. Genau diese Annahmen über die Gebäudekontur sollen in diesem Abschnitt in den Optimierungsprozess integriert werden.

Die modellbasierten Ansätzen zur Gebäudeextraktion (s. Abschnitt 6.1) bringen über das jeweilige Modell Annahmen über die Gebäudegeometrie in den Extraktionsprozess ein. Aufgrund der problematischen Definition geeigneter Objektmodelle wird im Rahmen dieser Arbeit von einem modellbasierten Ansatz zur Ermittlung der Gebäudekontur abgesehen. Stattdessen wird die im letzten Abschnitt vorgestellte Adaption an Bildkanten mit Nebenbedingungen („Constraints“) versehen. Anhand dieser Nebenbedingungen soll eine gebäudespezifische Geometrie in den Optimierungsprozess

einfließen.

#### 6.4.1 Einführen gebäudespezifischer Nebenbedingungen

Im vorhergehenden Abschnitt wurde der Adaptionprozess als ein Optimierungsproblem beschrieben. Der einfachste Weg bei einer Optimierung Nebenbedingungen zu berücksichtigen ist, zu der zu optimierenden Zielfunktion Strafterme hinzuzufügen, um unerwünschte Konstellationen bei dem Optimierungsprozess zu benachteiligen. Entfernt sich eine Lösung von der Randbedingung, werden die Strafterme sukzessive erhöht, bis die Randbedingung in hinreichender Weise eingehalten wird. Dadurch wird ein beschränktes Optimierungsproblem in ein unbeschränktes überführt. Jedoch führen diese „weichen“ Kriterien, insbesondere bei einer größeren Zahl an Nebenbedingungen, zu einer schlechten Konvergenz (FLETCHER 1987): Es ist möglich, dass beim Optimieren nur die Nebenbedingungen, nicht aber die eigentliche Zielfunktion berücksichtigt werden. Weiterhin ist es bei unterschiedlichen Nebenbedingungen schwer, diese adäquat zu gewichten.

Eine Optimierung der Polygongüte unter noch zu definierenden gebäudespezifischen Nebenbedingungen  $c$  lässt sich wie folgt formalisieren:

$$\text{Maximiere } Q^{(V)} \quad \text{unter der Bedingung } c(V) = 0. \quad (6.10)$$

Eine dieser Art beschränkte Optimierung sei als „hart“ bezeichnet — in Abgrenzung zu einer „weichen“ Addition von Straftermen.

Die Standardmethoden zur Optimierung mit harten Nebenbedingungen weisen die erwähnten Probleme bei Verwendung von Straftermen nicht auf. Doch zum einen müssen große lineare Gleichungssysteme gelöst werden; zum anderen bleiben die guten Konvergenzeigenschaften Aktiver Konturen meist nicht gewahrt. Das in Abschnitt 6.4.2 vorgestellte Vorgehen zur Optimierung der Polygongüte unter Beachtung harter Nebenbedingungen weist diese Probleme nicht auf. Die Umsetzung von  $r$  Nebenbedingungen ist durch die Lösung eines linearen Gleichungssystems der Größe  $r \times r$  implementiert. Zudem ist die Einhaltung der Nebenbedingungen in den iterativen Ablauf zur Polygonadaptation integriert, was eine minimale Beeinträchtigung der Konvergenzeigenschaften gewährleistet.

#### Auswahl geeigneter Vertices

Die nun folgenden Nebenbedingungen für Gebäudekonturen kommen nur an ausgewählten Polygonstützpunkten zum Tragen. Die Erfahrung in der Anwendung der Polygonoptimierung mit Nebenbedingungen zeigt, dass Constraints mit Bedacht verwendet werden sollten. Sie sollten nur an wirklich geeigneten Vertices angewendet werden, da sonst die Bildung von Fehlkonstellationen begünstigt wird. In unsicheren Fällen ist es besser, einen Vertex nicht zu beschränken — dann erfolgt lediglich eine Optimierung gemäß des Bildgradienten. Aus diesem Grund unterliegt die Auswahl der zu beschränkenden Vertices mehreren Kriterien.

Ein wichtiges Kriterium ist der von zwei benachbarten Polygonsegmenten eingeschlossene Winkel. Dieser Winkel ergibt sich für den Stützpunkt  $v_i$  aus der Definition des Skalarprodukts:

$$\overline{v_{i-1}v_i} \cdot \overline{v_i v_{i+1}} = |\overline{v_{i-1}v_i}| |\overline{v_i v_{i+1}}| \cos \rho_i \quad (6.11)$$

$$\cos \rho_i = \frac{(x_i - x_{i-1})(x_{i+1} - x_i) + (y_i - y_{i-1})(y_{i+1} - y_i)}{\sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}} \quad (6.12)$$

Bei beiden anschließend folgenden Nebenbedingungen wird vorausgesetzt, dass die zu beschränkenden Vertices einen Winkel von etwa  $90^\circ$  einschließen; für sie gilt also  $|\rho_i - \pi/2| < \rho_{\max}$ . Die Menge der Vertices, die diesem Rechtwinkligkeitskriterium genügen, sei mit  $\mathcal{R}_{\text{all}}$  bezeichnet:

$$\mathcal{R}_{\text{all}} = \{v_i \mid |\rho_i - \pi/2| < \rho_{\max}\} \quad (6.13)$$

Die Zuordnung als rechter Winkel hat sich bei Vertices, die von zwei ebenfalls als nahezu rechtwinklig eingeordneten Stützpunkten umgeben sind, als besonders zuverlässig erwiesen. Deshalb werden aus  $\mathcal{R}_{\text{all}}$  nur die Vertices ausgewählt, die dieser Anforderung genügen:

$$\mathcal{R}_{\text{inner}} = \{v_i \mid (v_{i-1} \in \mathcal{R}_{\text{all}}) \wedge (v_{i+1} \in \mathcal{R}_{\text{all}})\} \quad (6.14)$$

Vertices, die zwar einen nahezu rechten Winkel aufweisen, aber den Abschluss einer Kette benachbarter rechtwinkliger Stützpunkte bilden, sind also nicht mehr in  $\mathcal{R}_{\text{inner}}$  enthalten.

### Rechtwinkligkeit

Wie schon erwähnt, sind rechte Winkel ein typisches Merkmal von Gebäuden. Aus diesem Grund wird für Vertices aus  $\mathcal{R}_{\text{inner}}$  eine Rechtwinkligkeits-Nebenbedingung eingeführt. Dadurch soll der rechte Winkel für diese Polygonsegmente forciert werden. Wie bei der Auswahl der zu beschränken Vertices geschieht auch die Beschränkung über das Skalarprodukt. Ein rechter Winkel innerhalb der Stützpunkte  $v_{i-1}$ ,  $v_i$  und  $v_{i+1}$  führt dazu, dass das Skalarprodukt der beiden dadurch aufgespannten Vektoren Null ist:

$$\overline{v_{i-1}v_i} \cdot \overline{v_i v_{i+1}} \stackrel{!}{=} 0 \quad (6.15)$$

Dadurch ergibt sich die Nebenbedingung für einen rechten Winkel:

$$c^{(r)} = (x_i - x_{i-1})(x_{i+1} - x_i) + (y_i - y_{i-1})(y_{i+1} - y_i) \quad (6.16)$$

### Gemeinsame Orientierung der Polygonsegmente

Ein weiteres typisches Merkmal von Gebäudeumrissen ist deren regelmäßige Gestalt. Gebäude weisen meist eine Grundorientierung auf, zu der die Konturlinien senkrecht oder parallel verlaufen. Ge-

nau diese Grundorientierung soll über Nebenbedingungen geeigneten Polygonsegmenten auferlegt werden. Sowohl bei der Bestimmung der Grundorientierung wie auch bei deren Forcierung werden nur Vertices betrachtet, die einen nahezu rechten Winkel einschließen ( $\mathcal{R}_{\text{inner}}$ ). Für diese Stützpunkte kommt auch die bereits erwähnte Rechtwinkligkeits-Nebenbedingung zum Tragen. Rechten Winkeln wird eine besondere Wichtigkeit eingeräumt, da sie als essentieller Bestandteil des Gebäudes betrachtet werden.

Die Grundorientierung des Gebäudepolygons wird wie folgt bestimmt:

$$\bar{\varrho} = \frac{1}{|\mathcal{R}_{\text{inner}}|} \sum_{i \in \mathcal{R}_{\text{inner}}} \varrho_i \bmod \pi/2 \quad (6.17)$$

Anschließend wird die maximale absolute Abweichung der Orientierungen der Vertices in  $\mathcal{R}_{\text{inner}}$  bestimmt:

$$\tilde{i} = \operatorname{argmax}_{i \in \mathcal{R}_{\text{inner}}} |\bar{\varrho} - (\varrho_i \bmod \pi/2)| \quad (6.18)$$

Ist die Abweichung größer als eine vorgegebene maximal zulässige Winkeldifferenz, so wird  $v_{\tilde{i}}$  aus  $\mathcal{R}_{\text{inner}}$  entfernt. Die Berechnung von  $\bar{\varrho}$  und die anschließende Prüfung erfolgt nun erneut. Im Bedarfsfall wird dies iteriert bis die Abweichung unterhalb der Schwelle liegt. Die verbleibende Menge zu beschränkender Vertices sei mit  $\mathcal{R}'_{\text{inner}}$  bezeichnet. Dies sorgt dafür, dass rechte Winkel, die anders angeordnet sind als die Mehrheit der rechten Winkel des Polygons, nicht in den Mittelungsprozess einfließen. Andernfalls würde  $\bar{\varrho}$  durch solche Vertices zu sehr verfälscht werden.

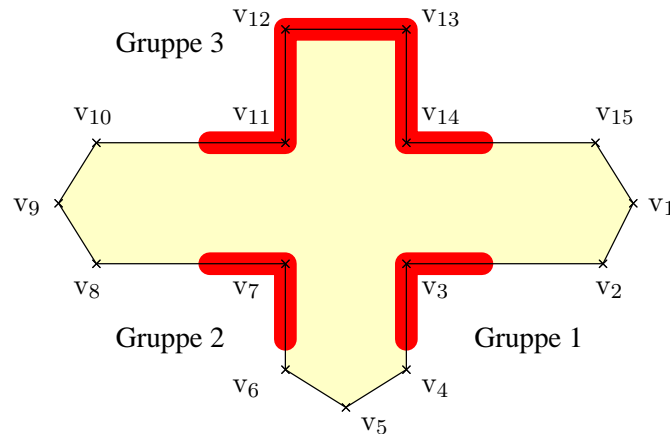
Die Grundorientierung wird nun den Vertices in der möglicherweise reduzierten Menge  $\mathcal{R}_{\text{inner}}$  zugeschrieben. Für diese Stützpunkte kommt immer die Rechtwinkligkeits-Nebenbedingung zum Tragen. Folgen mehrere rechtwinklig orientierte Vertices aufeinander, so genügt es, einen dieser Vertices bezüglich seiner Orientierung zu beschränken. Da die anderen benachbarten rechtwinkligen Vertices bei Einhaltung der Rechtwinkligkeits-Bedingung „starr“ mit dem ausgewählten Vertex verbunden sind, wird auch ihre Orientierung entsprechend  $\bar{\varrho}$  korrigiert. Aus diesem Grund werden Gruppen aufeinanderfolgender (nahezu) rechtwinkliger Vertices gebildet. Dies ist in Abbildung 6.7 veranschaulicht. Für jede dieser Gruppen wird für die Umsetzung der Orientierungs-Nebenbedingung ein Polygonsegment  $\overline{v_i v_{i+1}}$  bestimmt, das senkrecht zur Grundorientierung ausgerichtet sein muss. Wie bei dem Rechtwinkligkeits-Constraint in Gleichung 6.16, wird dies über das Skalarprodukt mit einem (Einheits-)Vektor in Grundorientierung implementiert:

$$c^{(o)} = x^\perp (x_{i+1} - x_i) + y^\perp (y_{i+1} - y_i). \quad (6.19)$$

Dabei bezeichnet  $v^\perp$  den durch  $\bar{\varrho}$  definierten Normalenvektor :

$$v^\perp = (x^\perp, y^\perp) = \frac{1}{\sqrt{1 + \tan^2 \bar{\varrho}}} (1, \tan \bar{\varrho}). \quad (6.20)$$





**Abbildung 6.7:** Schematische Darstellung zur Gruppierung in aufeinander folgende Polygonvertices mit nahezu rechtem Winkel.  $\mathcal{R}_{\text{inner}}$  besteht aus den Stützpunkten  $v_5, v_6, v_{12}, v_{13}, v_{18}, v_{19}, v_{20}$  und  $v_{21}$ . Daraus resultieren die drei rot unterlegten Vertexgruppen.

### 6.4.2 Umsetzung der Nebenbedingungen

Speziell zum Zusammenspiel mit der Optimierung Aktiver Konturen wird von FUA UND BRECHBÜLER (1996) ein Vorgehen für das Einhalten von Nebenbedingungen vorgestellt, das auf der Arbeit von BRECHBÜLER (1995) aufbaut. Nach diesem Ansatz werden die bereits eingeführten gebäudespezifischen Nebenbedingungen implementiert. Der eigentliche Adaptionsprozess und die Umsetzung der Nebenbedingungen werden dabei entkoppelt. Der durch Nebenbedingungen eingeschränkte Optimierungsprozess besteht aus zwei wesentlichen Schritten:

1. Die Nebenbedingungen werden durch die Projektion der aktuellen Polygonstützstellen  $V$  auf die „Fläche“, die durch die Nebenbedingungen aufgespannt wird, verwirklicht. Die Lösung der nichtlinearen Constraint-Gleichungen erfolgt durch Linearisieren und Newton-Schritte.
2. Maximierung der Zielfunktion  $Q^{(V)}$ . Dies geschieht durch Projektion ihres Gradienten tangential zur Constraint-Fläche. Die Suche nach dem Maximum erfolgt in Richtung der Projektion, so dass der resultierende Zustandsvektor sich nicht zu weit von der Constraint-Fläche entfernt.

Auf diese beiden Schritte wird nachfolgend näher eingegangen. Voranzustellen ist ein Implementierungsaspekt: Da die Polygonvertices zweidimensionale Punkte sind, gibt es bei einer Translation für jeden Punkt auch zwei Freiheitsgrade. Ein Polygon mit  $n$  Stützpunkten hat also  $2n$  Freiheitsgrade. Implementiert wird der Zustandsvektor eines Polygons als

$$\mathbf{v} = (x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n) = (v_1, v_2, \dots, v_{2n}). \quad (6.21)$$

Falls im Folgenden von  $2n$  Variablen für die Polygonvertices die Rede ist, sind also die beiden Freiheitsgrade je Vertex die Ursache.

### Projektion auf die Constraint-Fläche

Zuerst wird die Polygonverschiebung  $\Delta \mathbf{v}$  gesucht, die zu einer annähernden Einhaltung der Nebenbedingungen  $\mathbf{c}$  führt, also  $\mathbf{c}(\mathbf{v} + \Delta \mathbf{v}) \approx 0$ . Bei kleinen Verschiebungen  $\Delta \mathbf{v}$ , lassen sich die nichtlinearen Nebenbedingungen durch eine Taylorreihe erster Ordnung linear approximieren:

$$\mathbf{c}(\mathbf{v} + \Delta \mathbf{v}) \approx \mathbf{c}(\mathbf{v}) + A^T \Delta \mathbf{v} \quad (6.22)$$

Dabei ist  $A$  die Jakobi-Matrix der Nebenbedingungen. Bei  $n$  Polygonstützpunkten und  $r$  Nebenbedingungen hat  $A$  die Form

$$A = \frac{\partial \mathbf{c}}{\partial \mathbf{v}} = \nabla \mathbf{c}^T = \begin{pmatrix} \frac{\partial c_1}{\partial v_1} & \frac{\partial c_2}{\partial v_1} & \dots & \frac{\partial c_r}{\partial v_1} \\ \frac{\partial c_1}{\partial v_2} & \frac{\partial c_2}{\partial v_2} & \dots & \frac{\partial c_r}{\partial v_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial c_1}{\partial v_n} & \frac{\partial c_2}{\partial v_n} & \dots & \frac{\partial c_r}{\partial v_n} \end{pmatrix}. \quad (6.23)$$

Um am Vertex  $v_i$  einen rechten Winkel zu forcieren, wird eine Rechtwinkligkeits-Nebenbedingung,  $c^{(r)}$ , nach Gleichung 6.16 definiert. Deren Ableitungen ergeben sich wie folgt:

$$\frac{\partial c^{(r)}}{\partial x_{i-1}} = x_i - x_{i+1} \quad (6.24)$$

$$\frac{\partial c^{(r)}}{\partial y_{i-1}} = y_i - y_{i+1} \quad (6.25)$$

$$\frac{\partial c^{(r)}}{\partial x_i} = x_{i+1} - 2x_i + x_{i-1} \quad (6.26)$$

$$\frac{\partial c^{(r)}}{\partial y_i} = y_{i+1} - 2y_i + y_{i-1} \quad (6.27)$$

$$\frac{\partial c^{(r)}}{\partial x_{i+1}} = x_i - x_{i-1} \quad (6.28)$$

$$\frac{\partial c^{(r)}}{\partial y_{i+1}} = y_i - y_{i-1} \quad (6.29)$$

Um die Orientierung von Polygonsegmenten zu vereinheitlichen, wird durch Gleichung 6.19 eine Nebenbedingung  $c^{(o)}$  definiert. Deren Ableitungen lauten:

$$\frac{\partial c^{(o)}}{\partial x_{i-1}} = 0 \quad (6.30)$$

$$\frac{\partial c^{(o)}}{\partial y_{i-1}} = 0 \quad (6.31)$$

$$\frac{\partial c^{(o)}}{\partial x_i} = -x_i^\perp \quad (6.32)$$

$$\frac{\partial c^{(o)}}{\partial y_i} = -y^\top \quad (6.33)$$

$$\frac{\partial c^{(o)}}{\partial x_{i+1}} = x^\top \quad (6.34)$$

$$\frac{\partial c^{(o)}}{\partial y_{i+1}} = y^\top \quad (6.35)$$

Die Nebenbedingungen werden durch die Anwendung von Newtonschritten (s. z. B. BRONSTEIN UND SEMENDJAJEW 1991, Kapitel 7.1.2.3) näherungsweise erfüllt. Dazu wird iterativ die Gleichung

$$A^T \Delta \mathbf{v} = -\mathbf{c} \quad (6.36)$$

nach  $\Delta \mathbf{v}$  gelöst. In jedem Schritt wird dabei der Zustandsvektor  $\mathbf{v}$  aktualisiert:

$$\mathbf{v} \leftarrow \mathbf{v} + \Delta \mathbf{v} \quad (6.37)$$

Da die Zahl der Nebenbedingungen üblicherweise geringer ist als  $2n$ , hat Gleichung 6.36 mehrere Lösungen. Davon führt zwar jede zu einem annähernden Einhalten der Constraints, doch mit unterschiedlicher Entfernung von der Ausgangskonstellation. Da der Zustandsvektor  $\mathbf{v}$  bereits als eine gute Lösung der zur optimierenden Zielfunktion angesehen wird, sollte die Verschiebung  $\Delta \mathbf{v}$  zur Einhaltung der Nebenbedingungen so klein wie möglich sein. Aus diesem Grund wird als Resultat die kürzeste Entfernung für  $\Delta \mathbf{v}$  gewählt. Durch diese Wahl wird  $\Delta \mathbf{v}$  auf den Spaltenraum von  $A$  begrenzt. Komponenten von  $\Delta \mathbf{v}$ , die orthogonal zum Spaltenraum sind, ändern zwar nichts an der Einhaltung der Nebenbedingungen, verlängern jedoch  $\Delta \mathbf{v}$ . Damit lässt sich  $\Delta \mathbf{v}$  auch zerlegen in

$$\Delta \mathbf{v} = A \mathbf{u} . \quad (6.38)$$

Die Bestimmung von  $\mathbf{u}$  erfolgt durch Lösen des quadratischen, positiv definiten Gleichungssystems

$$A^T \Delta \mathbf{v} = A^T A \mathbf{u} = -\mathbf{c} . \quad (6.39)$$

Mit  $\mathbf{u}$  erhält man durch Gleichung 6.38 auch eine Lösung für  $\Delta \mathbf{v}$ . Dadurch nähert man sich der Constraint-Fläche; iteratives Anwenden dieser Newtonschritte führt letztlich zum Einhalten der Nebenbedingungen.

Wie noch in Abschnitt 6.6 ausgeführt werden wird, ist eine exakte Einhaltung der Nebenbedingungen in jedem Optimierungsschritt nicht notwendig. In jeder Optimierungsiteration wird nur eine Projektion auf die Constraint-Fläche vollzogen. Zusammen mit der Zielfunktion sollen so auch die Nebenbedingungen konvergieren. Um zu gewährleisten, dass sich die Vertexpositionen nicht zu weit von der Constraint-Fläche entfernen und somit die Konvergenz gefährden, müssen die Adaptionschritte beschränkt werden. Auf dieses Thema wird im Folgenden eingegangen.

### Beschränkte Maximierung der Zielfunktion

Nach der Projektion des Zustandsvektors  $\mathbf{v}$  orthogonal zur Constraint-Fläche mit dem Ziel, die Nebenbedingungen umzusetzen, folgt nun ein Schritt zur Adaption des Polygons an Bildkanten, wie in Abschnitt 6.3 beschrieben. Die bei der Polygonadaption vollzogene Verschiebung sei mit  $\Delta\tilde{\mathbf{v}}$  bezeichnet und der resultierende Zustandsvektor mit  $\tilde{\mathbf{v}}$ . Der Adaptionsschritt ist völlig unabhängig von den Nebenbedingungen. Um zu gewährleisten, dass sich der neue Zustandsvektor durch den Adaptionsschritt nicht zu weit von der Constraint-Fläche entfernt, wird  $\Delta\tilde{\mathbf{v}}$  diesbezüglich beschränkt. Dabei wird die kleinstmögliche Entfernung von der Constraint-Fläche bei bestmöglicher Umsetzung von  $\Delta\tilde{\mathbf{v}}$  angestrebt. Bei den geringen Verschiebungen, wie sie im Adaptionsprozess üblich sind, kann dies durch die Projektion auf den Untervektorraum tangential zur Constraint-Fläche (Nullraum von  $A$ ) erreicht werden. Dazu wird die Verschiebung bei der Polygonadaption,  $\Delta\tilde{\mathbf{v}}$ , als Gradient der Zielfunktion  $Q^{(\mathbf{v})}$  betrachtet. Da  $\Delta\tilde{\mathbf{v}}$  auf die Richtungskomponente tangential zur Constraint-Fläche beschränkt werden soll, werden die Richtungskomponenten orthogonal zur Constraint-Fläche „entfernt“. Zur Ermittlung der orthogonalen Komponenten wird versucht, den Gradienten der Zielfunktion so gut wie möglich durch eine Linearkombination von Constraint-Normalen darzustellen, also die  $\lambda$  zu bestimmen, für die gilt

$$A\lambda = \Delta\tilde{\mathbf{v}}. \quad (6.40)$$

Da dieses Gleichungssystem überbestimmt ist, wird mit dem  $r \times r$  System

$$A^T A\lambda = A^T \Delta\tilde{\mathbf{v}} \quad (6.41)$$

die Lösung mit dem kleinsten quadratischen Fehler  $|\Delta\tilde{\mathbf{v}} - A\lambda|^2$  bestimmt.  $A\lambda$  ist dann die Richtungskomponente von  $\Delta\tilde{\mathbf{v}}$ , die orthogonal zur Constraint-Fläche ist. Durch Subtraktion dieses Normalenvektors wird der Polygonadaptionsschritt auf die Richtungskomponente tangential zur Constraint-Fläche reduziert:

$$\Delta\tilde{\mathbf{v}} \leftarrow \Delta\tilde{\mathbf{v}} - A\lambda \quad (6.42)$$

Die beschriebenen Schritte zur Einhaltung der Nebenbedingungen, einerseits die Projektion des Zustandsvektors auf die Constraint-Fläche, andererseits die Korrektur des Optimierungsschritts bezüglich der Constraint-Fläche, laufen somit in zwei zueinander senkrechten Unterräumen ab. Einmal im Zeilenraum von  $A$ , das andere Mal im Nullraum von  $A^T$ .

Die Lösung der Kleinste-Fehlerquadrate-Gleichungen 6.39 und 6.41 erfolgt über die  $r \times r$  Matrix  $A^T A$ . Da die Zahl der Nebenbedingungen,  $r$ , hier immer kleiner ist als die Anzahl der Polygonstützpunkte,  $n$ , bleibt auch  $A^T A$  klein. Um Probleme mit der Lösung des Gleichungssystems bei singulären oder schlecht konditionierten Matrizen zu umgehen, ist die Lösung über eine Singulärwertzerlegung von  $A^T A$  implementiert. Über Einzelheiten hierzu sei beispielsweise auf PRESS ET AL. (1992, Kapitel 2.6) verwiesen.

## 6.5 Dynamische Verwaltung der Stützstellen

Die Adaption des Umrisspolygons sowie die Einhaltung der Nebenbedingungen ist über kleine Verschiebungsschritte der Vertices realisiert. Das Zusammenspiel der Polygonstützpunkte wird dabei nicht betrachtet. Da die initialen Positionen der Stützstellen durch die in Abschnitt 6.2.1 besprochene Polygonvereinfachung relativ willkürlich entlang der Konturlinie verteilt sind, gilt es jetzt, sie möglichst sinnvoll anzuordnen. Ferner sollen bestimmte Polygonkonstellationen unterdrückt werden. So sollte sich die Konturlinie nie mit sich selbst kreuzen. Auf diese Aspekte wird in den folgenden Abschnitten eingegangen.

### 6.5.1 Löschen von Vertices

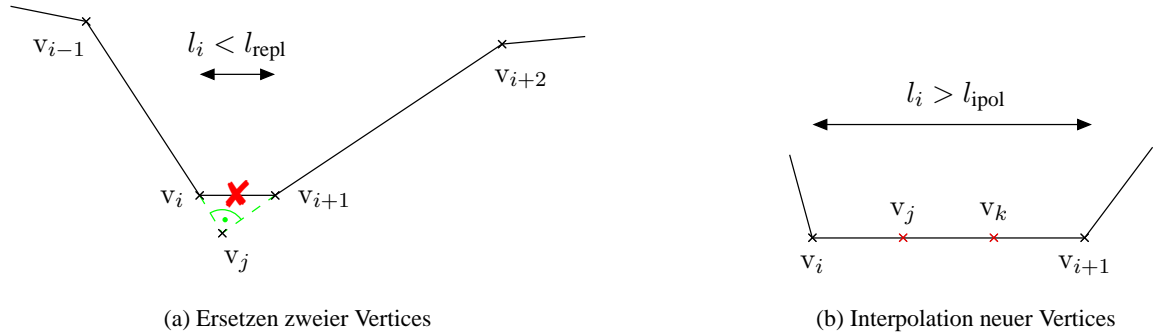
Immer dann, wenn drei aufeinanderfolgende Stützstellen kollinear zueinander sind, trägt der mittlere Punkt nicht zur Form des Polygonzugs bei. Deshalb werden Punkte, die nahezu kollinear zu ihren beiden Polygonnachbarn sind, gelöscht. Realisiert wird dies durch ein abermaliges Anwenden des in Abschnitt 6.2.1 vorgestellten Algorithmus zur Polygonvereinfachung. Außerdem wird dadurch immer eine möglichst einfache Beschreibung der Konturlinie, also eine möglichst geringe Anzahl an Stützstellen, erreicht.

Zudem wird das Polygon in jeder Iteration auf Konsistenz geprüft. Unsinnige Polygonkonstellationen sollen abgefangen und bereinigt werden. Im Fall von Polygonsegmenten, die von einem Vertex zu dessen Nachbarn und wieder zurück verlaufen, wird der doppelte Vertex wie auch der „Ausläufer“ entfernt. Kommt es bei einer der Polygoneformationen (Adaption/Nebenbedingung) zu einer Überschneidung des Polygons mit sich selbst, so wird dieser Deformationsschritt rückgängig gemacht.

Das Löschen unnötiger Vertices sowie die Konsistenzprüfung sind bei dem vorgestellten Vorgehen die entscheidenden Schritte zur Regularisierung des Polygons. Da bei der Polygonadaption ganz auf die bei Aktiven Konturen üblichen internen Energien verzichtet wird (s. auch Diskussion in Abschnitt 6.3), wird die Verteilung der Polygonvertices maßgeblich durch die Polygonvereinfachung kontrolliert.

### 6.5.2 Ausprägen rechter Winkel

Gelegentlich kommt es vor, dass zwei Polygonsegmente zwar fast rechtwinklig zueinander ausgerichtet, jedoch nicht direkt benachbart sind. Ein weiteres, kleines Segment liegt zwischen den beiden genannten Polygonsegmenten. Abbildung 6.8(a) verdeutlicht die geschilderte Polygonkonstellation. Trägt das kleine, mittlere Segment nicht wesentlich zur Güte des Polygons bei, wird es gelöscht, d. h. die beiden zugehörigen Vertices werden entfernt. Zuvor werden die beiden längeren Polygonsegmente, in Abbildung 6.8(a)  $\overline{v_{i-1}v_i}$  und  $\overline{v_{i+1}v_{i+2}}$ , bis zu ihrem gemeinsamen Schnittpunkt,  $v_j$ , verlängert.



**Abbildung 6.8:** Schematische Darstellung zum Einfügen neuer Polygonvertices. Falls zwei Polygonsegmente  $\overline{v_{i-1}v_i}$  und  $\overline{v_{i+1}v_{i+2}}$  übernächste Nachbarn sind und annähernd einen rechten Winkel formen, werden sie bis zu ihrem Schnittpunkt  $v_j$  verlängert (a). Bei langen Polygonsegmenten werden zwei neue Stützpunkte eingefügt (b).

Folgende Voraussetzung für das Ersetzen der Stützpunkte zugunsten einer rechtwinkligen Konstellation müssen gegeben sein:

Die beiden an das mittlere Segment  $\overline{v_i v_{i+1}}$  angrenzenden Segmente müssen eine Mindestlänge aufweisen. Das eingeschlossene Segment  $\overline{v_i v_{i+1}}$  darf nur relativ kurz sein, um große Änderungen in der Gestalt des Polygons zu vermeiden. Ferner darf die Kantengüte des mittleren Segments,  $q_{i,i+1}^{(V)}$ , einen gewissen Betrag nicht überschreiten. Dadurch soll ausgeschlossen werden, dass das kurze Segment  $\overline{v_i v_{i+1}}$  tatsächlich eine Kante der Hauskontur erfasst. Ist dies der Fall, so bleibt  $\overline{v_i v_{i+1}}$  aufgrund der hohen Kantengüte erhalten.

### 6.5.3 Einfügen neuer Vertices

Einerseits ist es das Ziel durch ein Polygon mit möglichst wenig Stützstellen eine möglichst einfache Beschreibung der Hauskontur zu erreichen. Andererseits ist es so schwer möglich, Einzelheiten der Gebäudekontur zu erfassen. Das Polygon kann nur dann Ausprägungen der Konturlinien erfassen, wenn sich in der Nähe auch Vertices befinden. Denn nur die Stützpunkte werden im Adaptionsprozess bewegt. Aus diesem Grund werden, wie in Abbildung 6.8(b) dargestellt, in jeder Iteration die Polygonsegmente mit einer Länge größer als ein Schwellwert  $l_{ipol}$  bestimmt. Auf diesen Segmenten werden jeweils äquidistant zwei neue Vertices eingefügt.

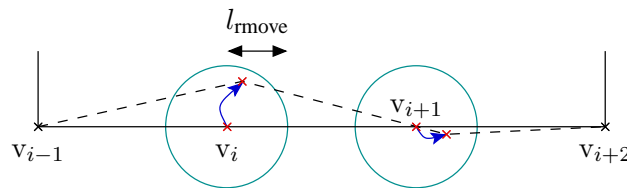
Durch Bewegen dieser beiden neuen Vertices können Abweichungen der Konturlinien von der geraden Verbindung von  $v_i$  nach  $v_{i+1}$  erfasst werden. Ändert sich die Position der beiden neuen Vertices nur geringfügig im Adaptionsprozess, wird keine neue Konturausprägung ermittelt. Durch den Polygonvereinfachungsschritt nach Abschnitt 6.5.1 werden die neuen Vertices dann wieder entfernt.

### 6.5.4 Verschieben von Vertices

Durch die Beschränkung der Polygonadaption auf eine Nachbarschaft um jeden Vertex wird der Adaptionsschritt sehr lokal bestimmt. Wie in Abschnitt 6.3.2 beschrieben, wird innerhalb dieser Nachbarschaft die Verschiebung ermittelt, die zur besten Polygongüte führt. Dabei ist es nicht möglich, zuerst einen etwas schlechter bewerteten Verschiebungsschritt einzuschlagen, um nachfolgend mit einem weiteren Schritt eine insgesamt bessere Konstellation zu erreichen.

Aus diesem Grund wird diese Möglichkeit „nachträglich“ eingeräumt: Ausgewählte Vertices werden in jeder Iteration zufällig geringfügig verschoben. Von dieser zufälligen Verschiebung betroffen sind nur Stützpunkte, die die Hauskontur noch nicht gut genug erfassen. Vertices mit hoher zugeordneter Kantengüte und solche, die einen rechten Winkel aufspannen, sind davon ausgeschlossen.

Die Verschiebung erfolgt zufällig in einem vorgegebenen Umkreis um jeden Vertex. Mit jeder Iteration wird der Kreisradius  $l_{\text{move}}$  verringert. Die Überlegung dabei ist, dass sich die Polygonoptimierung schrittweise stabilisiert und gegen eine endgültige Lösung konvergiert. Anfangs sind dazu große Schritte nötig; bereits nach wenigen Iterationen sollten nur noch feine Korrekturschritte erfolgen. Eine Verringerung der maximalen Entfernung für eine zufällige Verschiebung entspricht also diesen Konvergenzerwartungen. Abbildung 6.9 illustriert das Vorgehen zur Vertexverschiebung an einem Bildbeispiel.



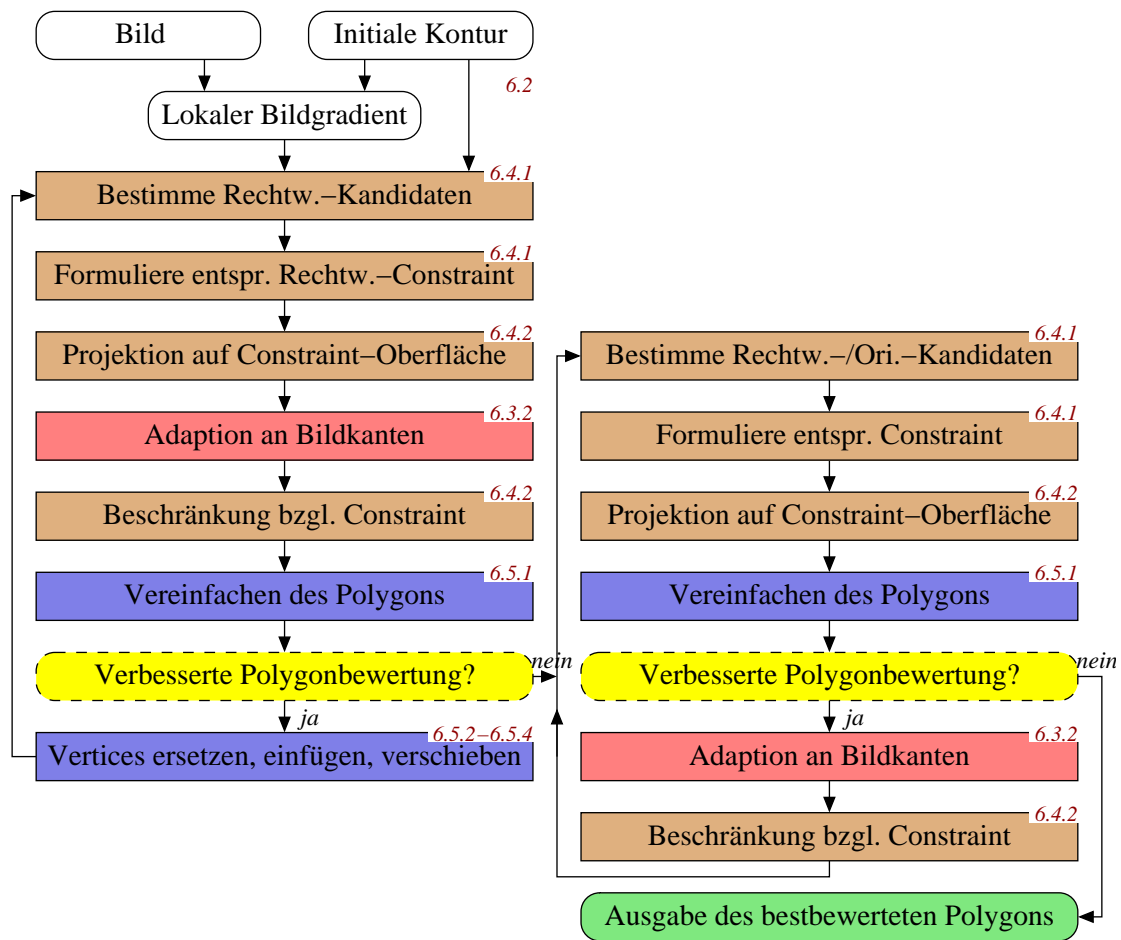
**Abbildung 6.9:** Schematische Darstellung zum zufälligen Verschieben schwacher Polygonvertices. Die schlecht bewerteten Stützpunkte  $v_i$  und  $v_{i+1}$  werden innerhalb des cyan dargestellten Umkreises mit Radius  $l_{\text{move}}$  zufällig verschoben. Die Verschiebung der beiden Vertices ist durch blaue Pfeile angedeutet.

## 6.6 Ablauf der Konturüberarbeitung

Für die in den vorhergehenden Abschnitten vorgestellten Teilschritte, die im Lauf der Konturüberarbeitung zum Einsatz kommen, wird das zeitliche Zusammenspiel im Gesamtablauf betrachtet. Dies ist in Abbildung 6.10 in Form eines Ablaufplans dargestellt. Ferner sind die Schritte in den Algorithmen 6 bis 8 als Pseudocode wiedergegeben.

Aus den Ausgangsdaten, einem Orthobild sowie einer initialen Gebäudekontur, werden in einem *Vorverarbeitungsschritt* gemäß Abschnitt 6.2 der lokale Bildgradient sowie eine vereinfachte initiale Kontur bestimmt (Algorithmus 6, Zeile 1–2<sup>1</sup>). Die darauf aufbauende Optimierungsphase unterteilt

<sup>1</sup>Mit „REFORM“ sei die Konvertierung der zweidimensionalen Polygonstützpunkte zu einem eindimensionalen Zustands-



**Abbildung 6.10:** Schematische Darstellung des Ablaufs zur Überarbeitung der Gebäudekonturen. Bei den jeweiligen Schritten wird auf die zugehörigen Abschnitte in diesem Kapitel verwiesen.

sich in zwei Durchläufe (Algorithmus 6, Zeile 3 und 4). In der *ersten Optimierungsschleife*, deren Ablauf in Algorithmus 7 wiedergegeben ist, adaptiert sich das Polygon schrittweise an den lokalen Bildgradienten; durch die in Abschnitt 6.4.1 definierte Rechtwinkligkeits-Nebenbedingung wird an geeigneten Vertices ein rechter Winkel forciert. Nach Bestimmung der zu beschränkenden Vertices (Algorithmus 7, Zeile 4) wird die Verschiebung  $\Delta \mathbf{v}$  ermittelt, die zur Projektion des Zustandsvektors auf die Constraint-Oberfläche nötig ist (Zeile 7). Ist der maximale Betrag aller Vertexverschiebungen  $\Delta \mathbf{v}$  kleiner als eine vorgegebene Schwelle, so wird die Verschiebung umgesetzt (Zeilen 8–10). Anschließend folgt in Zeile 11 eine unbeschränkte Adaption an den lokalen Bildgradienten, implementiert entweder als lokale Optimierung oder als globale Optimierung mit Dynamischer Programmierung. Wurde die Verschiebung zur Einhaltung der Nebenbedingung,  $\Delta \mathbf{v}$ , akzeptiert, so wird jetzt der unbeschränkte Adaptionsschritt bezüglich der Constraintfläche beschränkt (Zeile 13). Die

vektor bzw. umgekehrt bezeichnet.



Anzahl der Polygonstützpunkte wird durch die in Abschnitt 6.5 beschriebene dynamische Stützpunktverwaltung ständig variiert, um noch nicht erfasste Konturausprägungen im Adaptionsprozess berücksichtigen zu können (die Zeilen 15 sowie 21–23 entsprechen den in den Abschnitten 6.5.1–6.5.4 geschilderten Schritten). In jeder Iteration wird das Resultat der Konturüberarbeitung mit dem bisher besten Polygon verglichen (Zeile 16). Kriterium dafür sind die Anzahl der Rechtwinkligkeitskandidaten, deren Abweichung vom rechten Winkel, sowie die Güte des gesamten Polygons,  $Q^{(V)}$ , gemäß Gleichung 6.6 (Abschnitt 6.3.1). Aus den ersten beiden Angaben wird ein gemeinsames Maß abgeleitet:

$$Q_{\text{rect}}^{(V)} = |\mathcal{R}_{\text{inner}}| - \sum_{\{i \mid v_i \in \mathcal{R}_{\text{inner}}\}} \frac{|\rho_i - \pi/2|}{\rho_{\text{max}}} \quad (6.43)$$

Erhöht sich  $Q_{\text{rect}}^{(V)}$  gegenüber dem bisher bestbewerteten Polygon, so gilt das neue Polygon als das bessere und ersetzt dieses. Bei einer geringfügigen Verschlechterung von  $Q_{\text{rect}}^{(V)}$  und einer Verbesserung von  $Q^{(V)}$  gilt dies ebenso. Andernfalls ist keine Verbesserung durch den Überarbeitungsschritt eingetreten.

**Eingabe:** Konturpolygon  $V$ , Bild  $F$

**Ausgabe:** Optimiertes Konturpolygon  $V_{\text{opt}}$

POLYGONOPTIMIZATION( $V, F$ )

- (1)  $\mathbf{v} \leftarrow \text{SIMPLIFY}(\text{REFORM}(V))$
- (2)  $(|\nabla|^{(\text{loc})}, \varphi) \leftarrow \text{LOCALGRADIENT}(\mathbf{v}, F)$
- (3)  $\mathbf{v} \leftarrow \text{CONSTRAINEDADAPT1}(\mathbf{v}, |\nabla|^{(\text{loc})}, \varphi)$
- (4)  $\mathbf{v} \leftarrow \text{CONSTRAINEDADAPT2}(\mathbf{v}, |\nabla|^{(\text{loc})}, \varphi)$
- (5) **return** REFORM( $\mathbf{v}$ )

**Algorithmus 6:** Konturüberarbeitung, Kommandierung der beiden Optimierungsschleifen.

In jeder Iteration wird das Vorgänger-Polygon als Ausgangskontur herangezogen. Nach einer wählbaren Anzahl an Iterationen ohne Verbesserung (Standard: drei) wird die erste Optimierungsschleife abgebrochen. Das beste darin erzielte Gebäudepolygon wird Ausgangskontur für den zweiten Optimierungsdurchlauf.

Um den Optimierungsvorgang zeitlich zu begrenzen, wird die erste Optimierungsschleife auch im Fall eines stetig verbesserten Gütekriteriums nach einer wählbaren Anzahl an Iterationen abgebrochen. Ein Übergang zum zweiten Durchlauf erfolgt ebenfalls, sobald sich die Polygonfläche nicht mehr wesentlich ändert.

Bei Beendigung der ersten Optimierungsschleife wird davon ausgegangen, dass die wesentlichen Konturausprägungen erfasst wurden. Somit ist eine dynamische Stützpunktverwaltung für das weitere Vorgehen entbehrlich. Die *zweite Optimierungsschleife*, deren Ablauf in Algorithmus 8 wiedergegeben ist, kann als Feinabstimmung des Polygons verstanden werden. Neben dem Rechtwinkligkeits-Constraint für Vertices aus  $\mathcal{R}_{\text{inner}}$  kommt nun zusätzlich für die Vertices aus  $\mathcal{R}'_{\text{inner}}$  das Orientierungs-

**Eingabe:** Zustandsvektor  $\mathbf{v}$ , lokaler Bildgradient  $(|\nabla|^{(\text{loc})}, \varphi)$   
**Ausgabe:** Beschränkt optimierter Polygonzustandsvektor  $\mathbf{v}_{\text{opt}}$   
 CONSTRAINEDOPT1( $\mathbf{v}$ ,  $|\nabla|^{(\text{loc})}$ ,  $\varphi$ )

```

(1)   iter  $\leftarrow$  0; worse  $\leftarrow$  0
(2)    $\mathbf{v}_{\text{opt}} \leftarrow \mathbf{v}$ 
(3)   repeat
(4)     Bestimme  $\mathcal{R}_{\text{inner}}$ 
(5)     constrain  $\leftarrow$  false
(6)     if  $\mathcal{R}_{\text{inner}} \neq \emptyset$ 
(7)       Bestimme  $\Delta \mathbf{v}$  mit  $c^{(r)}(\mathbf{v} + \Delta \mathbf{v}) \approx 0$ 
(8)       if  $\Delta \mathbf{v} < \Delta \mathbf{v}_{\text{max}}$ 
(9)         constrain  $\leftarrow$  true
(10)       $\mathbf{v} \leftarrow \mathbf{v} + \Delta \mathbf{v}$ 
(11)      $\Delta \tilde{\mathbf{v}} \leftarrow \text{UNCONSTRAINEDADAPT}(\mathbf{v}, |\nabla|^{(\text{loc})}, \varphi)$ 
(12)     if constrain
(13)       Beschränkung des Adaptionsschritts:  $\Delta \tilde{\mathbf{v}} \leftarrow \Delta \tilde{\mathbf{v}} - A\lambda$ 
(14)     Realisation des neuen Zustandsvektors:  $\mathbf{v} \leftarrow \mathbf{v} + \Delta \tilde{\mathbf{v}}$ 
(15)      $\mathbf{v} \leftarrow \text{SIMPLIFY}(\mathbf{v})$ 
(16)     if ISBETTER1( $\mathbf{v}$ ,  $\mathbf{v}_{\text{opt}}$ )
(17)        $\mathbf{v}_{\text{opt}} \leftarrow \mathbf{v}$ 
(18)     else
(19)       worse  $\leftarrow$  worse + 1
(20)     iter  $\leftarrow$  iter + 1
(21)     RIGHTANGLESUBSTITUTE( $\mathbf{v}_{\text{opt}}$ )
(22)     INTERPOLATEVERTICES( $\mathbf{v}_{\text{opt}}$ )
(23)     RANDOMLYMOVEWEAKVERTICES( $\mathbf{v}_{\text{opt}}$ , iter)
(24)   until (iter > iter_max) or (worse > worse_max)
(25)   return  $\mathbf{v}_{\text{opt}}$ 
```

**Algorithmus 7:** Konturüberarbeitung, erste Optimierungsschleife. Das Polygon wird unter der Nebenbedingung, rechte Winkel zu forcieren, an die Gebäudekanten adaptiert. Die Polygonstützstellen werden dynamisch verwaltet.

Constraint zum Tragen (Algorithmus 8, Zeilen 4, 7 und 13). Bei einer festen Anzahl an Stützpunkten<sup>2</sup> wird das Polygon so entsprechend der Hauptorientierung ausgerichtet. Üblicherweise sind die Veränderungen gering und deshalb auch nur wenige Schritte dafür erforderlich. Die Abbruchkriterien (Funktion „ISBETTER2()“, Zeile 16) werden so modifiziert, dass nun auch die Übereinstimmung mit der Hauptorientierung berücksichtigt wird. Die mittlere Abweichung von der Hauptorientierung  $\bar{\varrho}$  ergibt sich als

$$\Delta \bar{\varrho} = \frac{1}{|\mathcal{R}'_{\text{inner}}|} \sum_{\{i \mid \mathbf{v}_i \in \mathcal{R}'_{\text{inner}}\}} |\bar{\varrho} - (\varrho_i \bmod \pi/2)|. \quad (6.44)$$

<sup>2</sup>Da das Polygon in jeder Iteration vereinfacht wird, kann die Anzahl der Stützpunkte in seltenen Fällen auch abnehmen.

Damit wird dann ein Gütekriterium für den zweiten Optimierungsdurchlauf definiert:

$$Q_{\text{ori}}^{(V)} = \frac{Q_{\text{rect}}^{(V)}}{1 + \min(\Delta\bar{\varrho}, \Delta\bar{\varrho}_{\text{max}})} \quad (6.45)$$

Das Polygon gilt als verbessert, wenn sich  $Q_{\text{ori}}^{(V)}$  nur leicht verschlechtert und die Polygongüte  $Q^{(V)}$  entsprechend verbessert oder umgekehrt. Analog zum ersten Durchlauf erfolgt auch hier ein Schleifenabbruch nach einer bestimmten Anzahl an Iterationen, nach einer gewissen Anzahl an Verschlechterungen sowie bei zu geringer Flächenveränderung. Endgültiges Resultat ist das am besten bewertete Gebäudepolygon.

**Eingabe:** Zustandsvektor  $\mathbf{v}$ , lokaler Bildgradient  $(|\nabla|^{(\text{loc})}, \varphi)$   
**Ausgabe:** Beschränkt optimierter Polygonzustandsvektor  $\mathbf{v}_{\text{opt}}$   
 CONSTRAINEDOPT2( $\mathbf{v}$ ,  $|\nabla|^{(\text{loc})}, \varphi$ )

```

(1)  iter ← 0; worse ← 0
(2)   $\mathbf{v} \leftarrow \mathbf{v}_{\text{opt}}$ 
(3)  repeat
(4)    Bestimme  $\mathcal{R}_{\text{inner}}$  und  $\mathcal{R}'_{\text{inner}}$ 
(5)    constrain ← false
(6)    if  $\mathcal{R}_{\text{inner}} \neq \emptyset$  or  $\mathcal{R}'_{\text{inner}} \neq \emptyset$ 
(7)      Bestimme  $\Delta\mathbf{v}$  mit  $c^{(r,o)}(\mathbf{v} + \Delta\mathbf{v}) \approx 0$ 
(8)      if  $\Delta\mathbf{v} < \Delta\mathbf{v}_{\text{max}}$ 
(9)        constrain ← true
(10)      $\mathbf{v} \leftarrow \mathbf{v} + \Delta\mathbf{v}$ 
(11)     $\Delta\tilde{\mathbf{v}} \leftarrow \text{UNCONSTRAINEDADAPT}(\mathbf{v}, |\nabla|^{(\text{loc})}, \varphi)$ 
(12)    if constrain
(13)      Beschränkung des Adaptionsschritts:  $\Delta\tilde{\mathbf{v}} \leftarrow \Delta\tilde{\mathbf{v}} - A\lambda$ 
(14)      Realisation des neuen Zustandsvektors:  $\mathbf{v} \leftarrow \mathbf{v} + \Delta\tilde{\mathbf{v}}$ 
(15)       $\mathbf{v} \leftarrow \text{SIMPLIFY}(\mathbf{v})$ 
(16)      if ISBETTER2( $\mathbf{v}$ ,  $\mathbf{v}_{\text{opt}}$ )
(17)         $\mathbf{v}_{\text{opt}} \leftarrow \mathbf{v}$ 
(18)      else
(19)        worse ← worse + 1
(20)      iter ← iter + 1
(21)  until (iter > iter_max) or (worse > worse_max)
(22)  return  $\mathbf{v}_{\text{opt}}$ 
```

**Algorithmus 8:** Konturüberarbeitung, zweite Optimierungsschleife. Das Polygon wird unter der Nebenbedingung, rechte Winkel sowie eine einheitliche Ausrichtung zu forcieren, an die Gebäudekanten adaptiert.

Die wesentlichen Schritte werden zur Veranschaulichung nun durch ein konkretes Bildbeispiel illustriert. Grundlage ist ein kleiner Bildausschnitt mit 50 cm Pixelauflösung, auf dem ein L-förmig

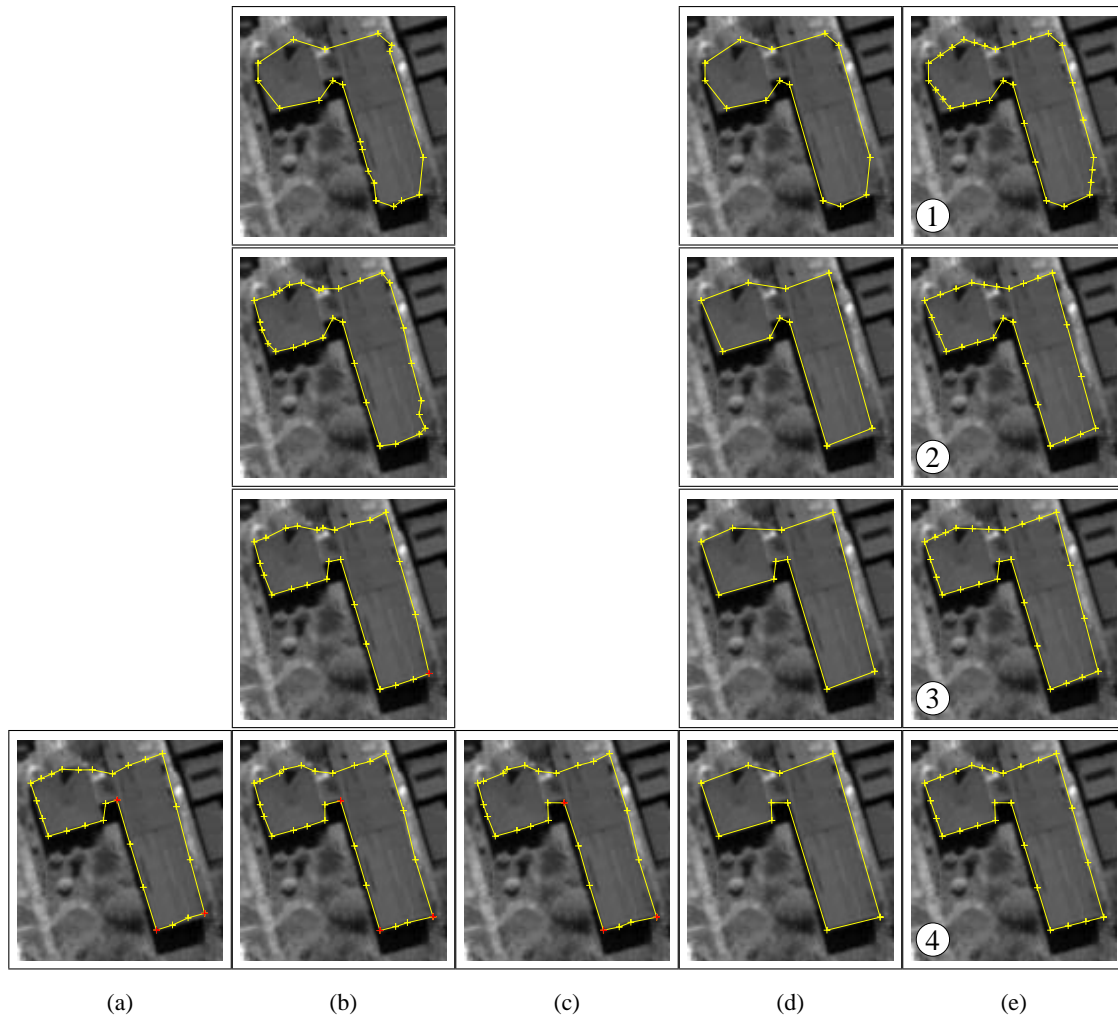
zusammengesetztes Gebäude zu sehen ist. Die initiale Kontur wurde durch Konturierung des nDOM abgeleitet (vgl. Abbildungen 6.2(a) und (b)). Die Abbildungen 6.11 und 6.12 zeigen die Zwischenergebnisse des ersten Optimierungsdurchlaufs. Jede Bildzeile (1–8) repräsentiert eine Iteration. Die Spalten geben die wesentlichen Arbeitsschritte wieder. Innerhalb der ersten drei Iterationen ist die Polygonoptimierung unbeschränkt, da die durch die Nebenbedingungen hervorgerufen Verschiebungen zu groß sind. In den nachfolgenden Iterationen kommen die Constraints zum Einsatz; die rot dargestellten Kreuze markieren die betroffenen Vertices. Im fünften Iterationsschritt wird das bestbewertete Polygon bestimmt. Die nachfolgenden drei Iterationen bewirken keine Verbesserung, deshalb wird die erste Optimierungsschleife abgebrochen. Abbildung 6.13 zeigt die Zwischenergebnisse des zweiten Optimierungsdurchlaufs. Bereits in der ersten Iteration wird das beste Resultat erzielt. Insgesamt wurden  $8 + 4 = 12$  Optimierungsiterationen benötigt. Die wesentlichen Gebäudekonturen und rechten Winkel wurden erfasst. An der oberen Gebäudemitte am Übergang von der einen Gebäudehälfte zur anderen liegt in Realität ebenfalls ein doppelter rechter Winkel vor. Doch erstens ist dieser Übergang im Bild nicht hinreichend genau erfasst. Zweitens wäre die kleine rechtwinklige Einbuchtung zu detailliert für die Genauigkeitsschwelle im Polygonvereinfachungsalgorithmus.

### 6.7 Evaluation der Konturüberarbeitung

Ziel dieses Abschnitts ist eine qualitative (Abschnitt 6.7.1) wie auch quantitative Bewertung des vorgeschlagenen Vorgehens zur beschränkten Polygonoptimierung. Wie in den vorhergehenden Kapiteln auch wird dazu einerseits die damit erreichbare Genauigkeit (Abschnitt 6.7.2), andererseits das Laufzeitverhalten untersucht (Abschnitt 6.7.3).

Zur Optimierung des Konturpolygons ohne Nebenbedingungen sind nach Abschnitt 6.3.2 zwei Strategien implementiert. Zum einen eine stützpunktweise lokale Optimierung, zum anderen eine globale Optimierung mittels dynamischer Programmierung. Aufgrund des wesentlich größeren Zustandsraums der globalen Optimierungsstrategie sind im Vergleich zur lokalen Optimierung deutlich längere Laufzeiten zu erwarten.

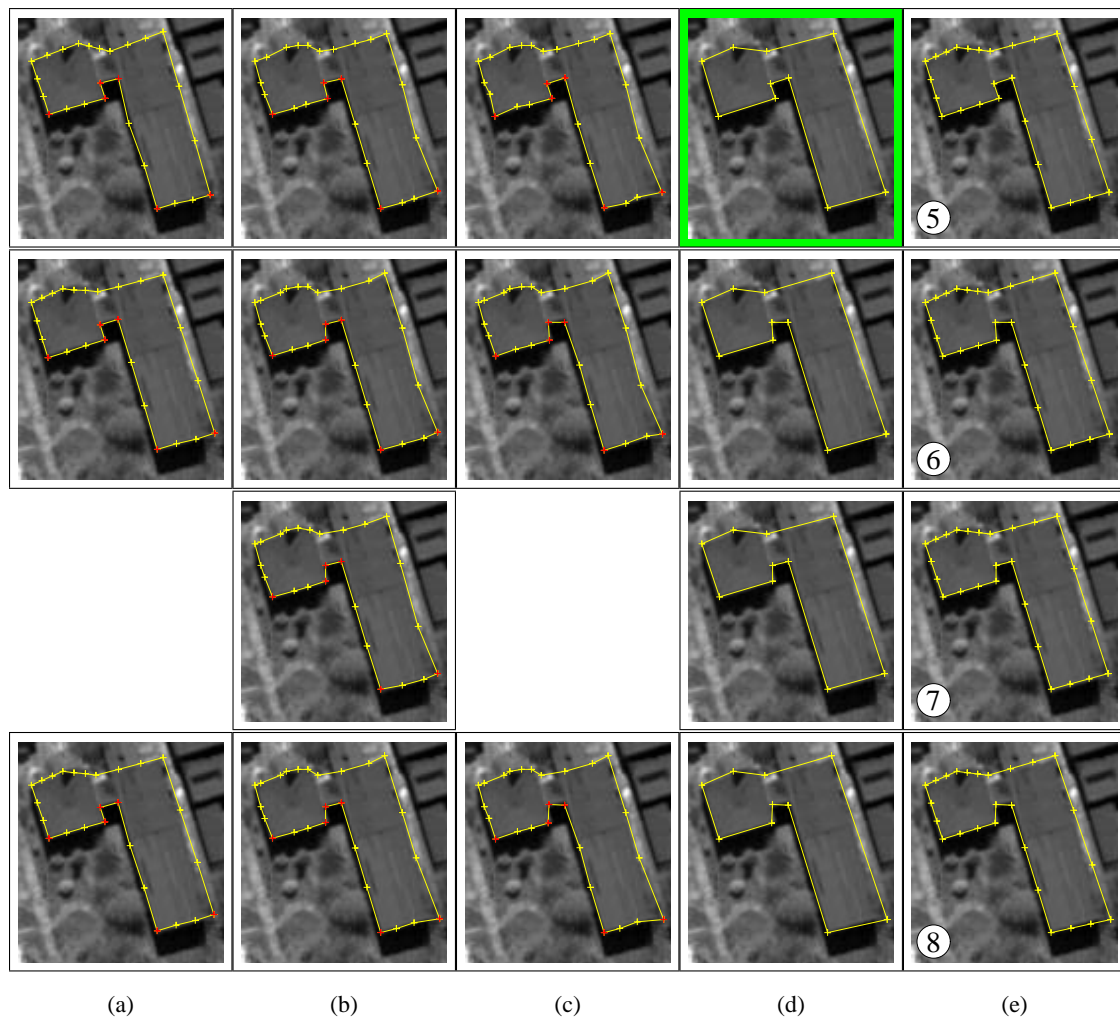
Wie in Abschnitt 6.6 dargestellt, gliedert sich der Ablauf der Konturüberarbeitung in zwei Phasen: Die erste Phase dient durch die ständige Variation der Anzahl der Stützpunkte der Erfassung der realen Konturgegebenheiten. In der zweiten Phase wird von einer nahezu korrekten Erfassung der Hauskontur ausgegangen. Hier finden lediglich feine Änderungen in der Ausrichtung der Polygonsegmente statt. In der ersten Optimierungsphase werden — bedingt durch das Einfügen neuer Vertices gemäß Abschnitt 6.5.3 — mehr Stützpunkte adaptiert als zur Beschreibung des Polygons nötig. Da in der zweiten Optimierungsphase immer mit der kleinstmöglichen Menge an Polygonstützpunkten gearbeitet wird, benötigt die eigentliche Adaption auch weniger Rechenzeit. Somit bietet sich auch eine Kombination aus lokaler und globaler Optimierung an: In der ersten Optimierungsschleife wird auf eine schnelle, lokale Optimierung zurückgegriffen, in der zweiten Schleife folgt



**Abbildung 6.11:** Konturüberarbeitung an einem Bildbeispiel — erste Optimierungsschleife (Teil 1). Die roten Kreuze markieren die Vertices, an denen die Rechtwinkligkeits-Nebenbedingung umgesetzt wird. Die festgehaltenen Schritte umfassen die Projektion auf die Constraint-Fläche (a), die Polygonadaption (b), die Beschränkung des Adaptionsschritts bzgl. der Nebenbedingung (c), die Polygonvereinfachung (d) sowie das Einfügen neuer Vertices (e).

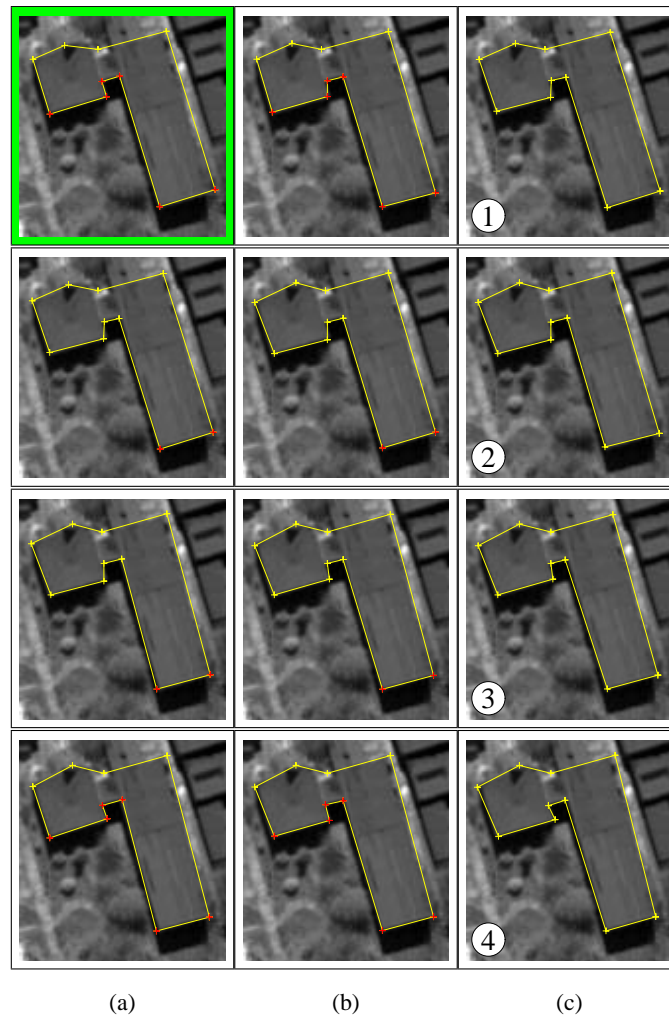
dann eine globale Optimierung. Dieser hybride Ansatz wird ebenfalls hinsichtlich der so erzielbaren Genauigkeit bewertet. Eine Laufzeitabschätzung ergibt sich durch Kombination der nachfolgend genannten Zeiten für die jeweils lokal oder global durchgeführte Adaption. Falls im Folgenden die Optimierungsstrategie nicht näher benannt sein sollte, wird aufgrund des Laufzeitverhaltens ein ausschließlich lokales Vorgehen verwendet.

Im Rahmen dieses Abschnitts wird auch betrachtet, inwieweit sich die Reihenfolge von Gebäudeunterteilung (Kapitel 5) und Konturüberarbeitung auf die Genauigkeit auswirkt. Einmal können größere Gebäudekomplexe in kleinere Einheiten unterteilt werden, deren Kontur anschließend sepa-



**Abbildung 6.12:** Konturüberarbeitung an einem Bildbeispiel — erste Optimierungsschleife (Teil 2). In der fünften Iteration wird das beste Resultat erzielt (grün umrandet). Die spaltenweise aufgeteilten Arbeitsschritte entsprechen denen in Abbildung 6.11.

rat überarbeitet wird. Anders herum können auch zuerst die Konturen der Gebäude überarbeitet und die größeren Gebäude dann aufgeteilt werden. Die beiden Vorgehensweisen führen zu unterschiedlichen Eingangsdaten für den jeweils zweiten Arbeitsschritt. Da die Orientierung der Trennlinien zur Unterteilung der Gebäudesegmente anhand der Segmentrandpixel bestimmt wird (s. Abschnitt 5.3), wirkt sich eine veränderte Konturlinie also darauf aus. Bei der Konturüberarbeitung gilt, dass kleinere Segmente im Allgemeinen zu weniger komplexen Konturlinien führen. Der Optimierungsprozess läuft an einem kleineren Gebäudeteil somit deutlich schneller ab als an einem größeren Gebäudekomplex. Ferner haben Experimente gezeigt, dass die Umsetzung von Nebenbedingungen bei sehr komplexen Gebäudepolygonen nicht mehr zu den gewünschten Resultaten führt. In solchen Fällen ist es beispielsweise schwierig, die Hauptorientierung des Gebäudepolygons zu bestimmen.

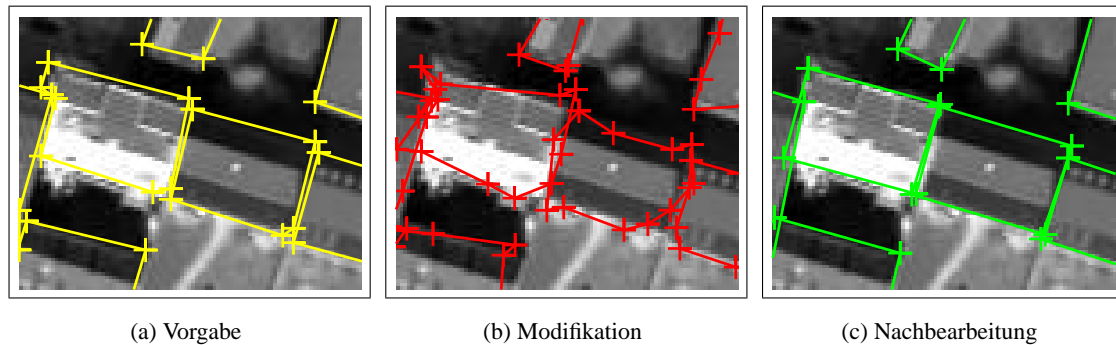


**Abbildung 6.13:** Konturüberarbeitung an einem Bildbeispiel — zweite Optimierungsschleife. Die roten Kreuze markieren die Vertices, an denen die Rechtwinkligkeits-Nebenbedingung umgesetzt wird. Zusätzlich kommt an ausgewählten Stützpunkten eine Nebenbedingung zur einheitlichen Ausrichtung der Polygonsegmente zum Tragen. Gleich in der ersten Iteration wird das beste Resultat erzielt (grün umrandet). Die spaltenweise aufgeteilten Arbeitsschritte entsprechen den drei ersten Schritten in Abbildung 6.11.

### 6.7.1 Qualitative Beurteilung

Die den folgenden Untersuchungen zugrunde liegenden Bilddaten stammen aus dem bereits in den beiden letzten Kapiteln verwendeten HRSC-A-Datensatz eines Berliner Stadtgebiets. Die Pixelauflösung beträgt 50 cm; aus dem Nadirkanal wird der lokale Bildgradient abgeleitet. Die Soll-Vorgabe der Gebäudepolygone wird analog zu den Auswertungen der beiden letzten Kapitel manuell aus den Bilddaten abgeleitet.

Da in diesem Kapitel die Verbesserung der Gebäudeumrisse im Vordergrund steht, gilt es, die



**Abbildung 6.14:** Künstliche Verschlechterung der Gebäudekonturvorgaben und anschließende Nachbearbeitung.

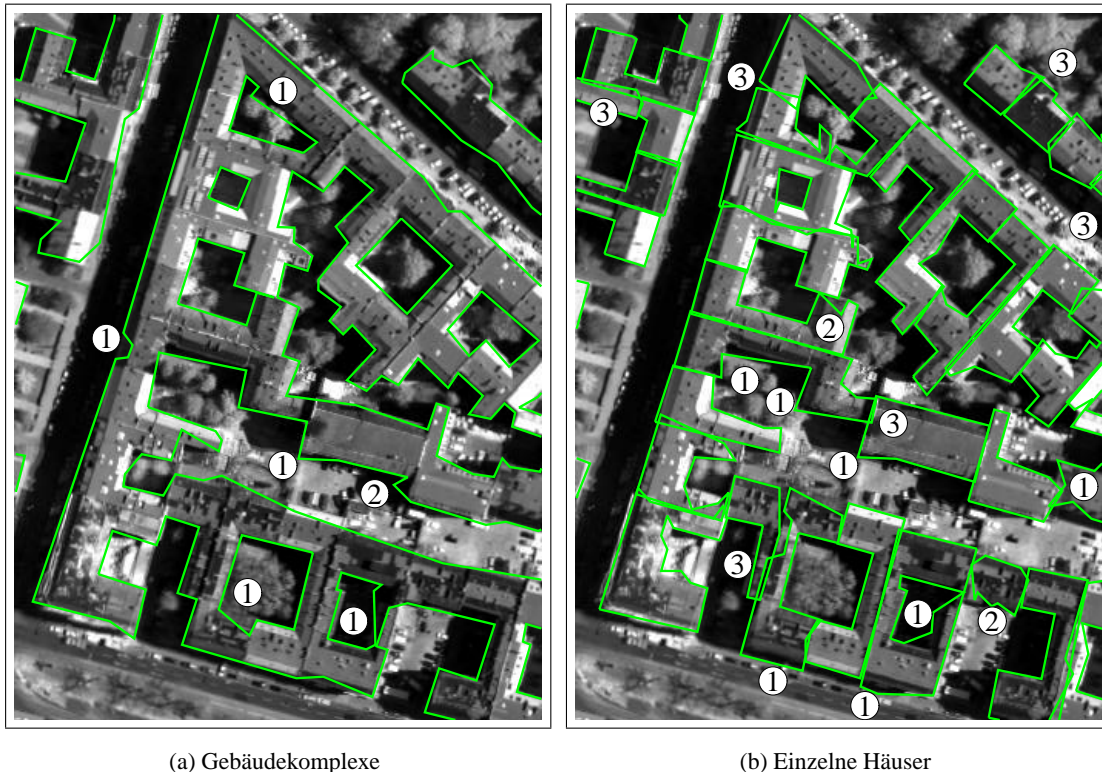
Konturlinien vor und nach der Überarbeitung zu vergleichen. Einmal werden die Ausgangskonturen durch eine künstliche Verschlechterung der Soll-Vorgabe bestimmt. Ein anderes Mal wird dafür das Ergebnis der Arbeitsschritte vorhergehender Kapitel herangezogen.

### Ein künstlich erzeugter Datensatz

Um unabhängig von den Resultaten einer vorhergehenden Gebäudesegmentierung Ausgangskonturen für die Konturüberarbeitung zu erhalten, wird auf die manuell bestimmten Konturvorgaben zurückgegriffen. Diese werden durch zufälliges Verschieben der Vertices wie folgt verschlechtert. Zuerst werden durch Interpolation neue Vertices eingefügt, so dass die Stützpunkte entlang der Konturlinie einen mittleren Abstand von 10 m aufweisen. Anschließend werden alle Vertices zufällig verschoben. Die Verschiebung ist normalverteilt um 0 m mit einer Varianz von 2 m ( $\sqrt{2}$  m in jeder Dimension). Abbildung 6.14 zeigt zur Veranschaulichung einen Bildausschnitt, der überlagert ist mit der Konturvorgabe bzw. mit deren Modifikation durch zufälliges Vertex-Verschieben sowie den überarbeiteten Konturen.

Zuerst wird der Fall betrachtet, dass aus den manuell bestimmten Konturvorgaben der *Gebäudekomplexe* ein Datensatz mit künstlich verschlechterten Gebäudeumrissen abgeleitet wird. Abbildung 6.15(a) zeigt einen Bildausschnitt überlagert mit dem Resultat der Konturüberarbeitung. Anhand einer solchen Überlagerung fallen dem Betrachter leicht Unstimmigkeiten zur vermuteten realen Kontur auf. Durch entsprechende Ziffern sind beispielhaft Stellen gekennzeichnet, an denen die überarbeiteten Gebäudepolygone merklich von den realen Umrissen abweichen. So liegt an den mit (1) markierten Stellen ein typisches Problem gradientenbasierter Adaptionansätze vor. Die Gradientenstärke der Gebäudekontur ist z. B. durch niedrigen Kontrast im Schatten gering. Zudem liegen in der Umgebung der realen Kontur stärker ausgeprägte Bildkanten vor, verursacht beispielsweise durch benachbarte Objekte, Dachelemente oder Schatten. Das Gebäudepolygon adaptiert sich in einem solchen Fall an die stärkere, allerdings falsche Bildkante.





**Abbildung 6.15:** Nachbearbeitung der aus der Vorgabe abgeleiteten Umrisspolygone (grün). Auf die nummerierten Fälle wird im Text eingegangen.

Durch die zufällige Verschiebung der Konturvorgaben kommt es bei wenigen Vertices vor, dass sie dann zu weit entfernt von dem realen Umriss sind. Die in Abschnitt 6.2.2 eingeführte Glättung der Gradientenstärke erzeugt einen die Bildkanten umgebenden „Anziehungsbereich“. Doch in dem mit (2) markierten Fall in Abbildung 6.15(a) liegt der betreffende Stützpunkt offensichtlich außerhalb dieses Bereichs in einer Region niedriger Gradientenstärke.

Werden nun künstlich verschlechterte Umrisse *einzelner Häuser* aus der Konturvorgabe abgeleitet, so resultieren die in Abbildung 6.15(b) dargestellten Konturlinien. Auffällig ist besonders, dass es teilweise zu Überschneidungen benachbarter Hauspolygone kommt (Beispiele sind mit der Ziffer (3) gekennzeichnet). Die Ursache dafür ist eine fehlende Fixierung der Trennlinien direkt aneinander grenzender Häuser während der Nachbearbeitung. Die Trennlinien werden für jedes der Hauspolygone separat überarbeitet. Verstärkt wird die „Gefahr“ der Überlappung durch die Art der Generierung der Ausgangspolygone: Die Konturen benachbarter Häuser werden unabhängig voneinander verschlechtert.

Des Weiteren sind in Abbildung 6.15(b) ebenfalls Adaptionen an falsche Bildkanten auszumachen (Ziffer (1)). Die mit (2) gekennzeichneten Stellen deuten auf Vertices, die bei der Überarbeitung nicht wesentlich verändert werden, obwohl sie die reale Kontur nicht erfassen. Grund hierfür ist die bereits

hohe Gradientenstärke an den Ausgangspunkten.

Dem Betrachter erscheinen die Hauspolygone nach dem Überarbeitungsschritt fast durchweg als verbessert. Dies ist vor allem der Polygonvereinfachung zuzuschreiben. Polygone mit möglichst langen geradlinigen Segmenten werden vom Betrachter eher als Gebäude eingestuft als zackige Konturlinien.

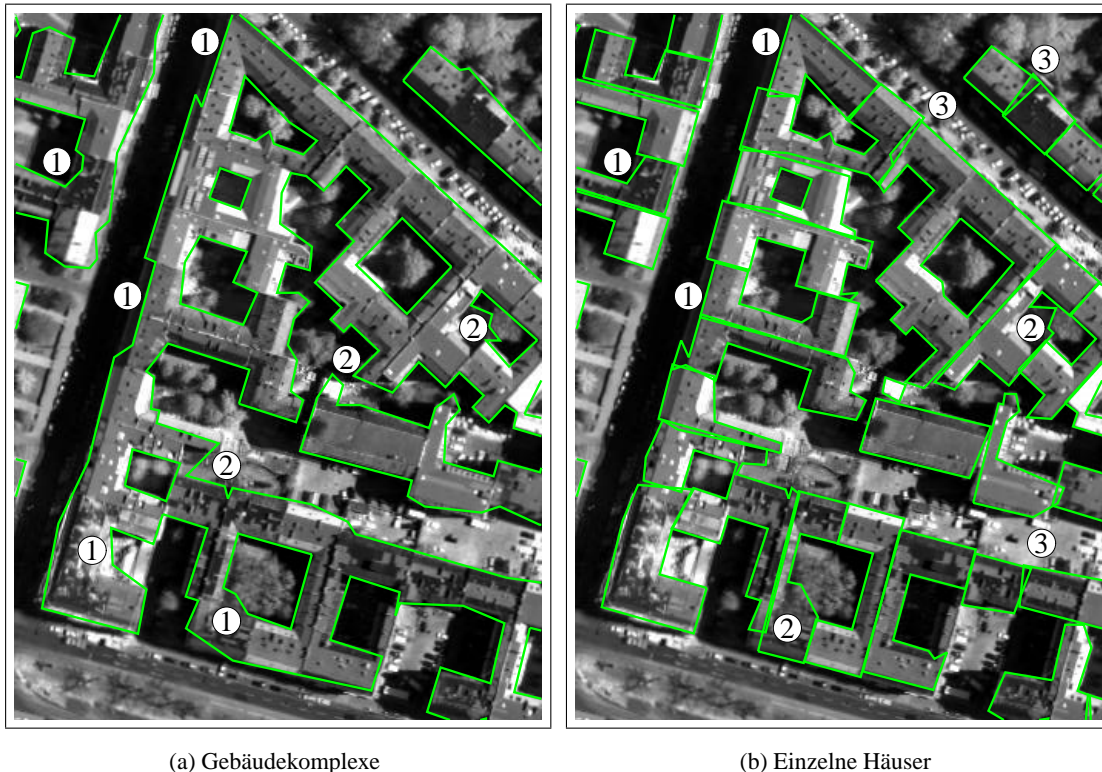
### Ein Datensatz aus vorhergehenden Arbeitsschritten

Im Unterschied zu den vorangegangenen Untersuchungen werden die Ausgangskonturen nun nicht künstlich erzeugt, sondern resultieren aus anderen Modulen dieser Arbeit. So werden die Ausgangspolygone einmal aus der gemäß Kapitel 4 durchgeführten Gebäudesegmentierung gewonnen. Diese Konturpolygone repräsentieren somit Gebäudekomplexe. Das andere Mal werden diese Gebäudesegmente nach Kapitel 5 in kleinere Einheiten unterteilt.

Abbildung 6.16(a) zeigt das Resultat der Konturüberarbeitung bei *Gebäudekomplexen*. Wie zuvor treten auch hier Fehladaptationen bei schwach ausgeprägten Gebäudekanten (1) sowie bei zu weit von der realen Gebäudekontur entfernt Ausgangspolygonen auf (2). Die schlechte Ausgangsposition ist auf eine ungenaue Gebäudesegmentierung zurückzuführen. Teilweise sind Innenhöfe nicht vollständig segmentiert, an anderen Stellen führen Überstrahlungen zu einer fehlerhaften Gebäude- bzw. Vegetationssegmentierung.

Betrachtet man die Konturüberarbeitung bei den in kleinere Einheiten *unterteilten Gebäudekomplexen*, so resultieren die im Bildausschnitt von Abbildung 6.16(b) überlagerten Polygone. Auch hier kommt es zu den zuvor genannten Fehladaptationen. Da die Überarbeitung benachbarter Häuser separat erfolgt, kann es zu Überschneidungen der Konturpolygone nach dem Nachbearbeitung kommen (3). Eine mögliche Lösung, um dies zu verhindern, wäre eine bereits erwähnte Fixierung der Trennlinien während der Überarbeitung. Allerdings sind in diesem realen Beispiel die Überschneidungen weniger stark ausgeprägt als im Datensatz mit den zufällig verschobenen Konturpolygonen. Hier ist das Polygonsegment, das die Trennlinie darstellt, für beide aneinander grenzenden Häuser dasselbe. Überlappungen sind in diesem Fall ausschließlich durch die separat durchgeführten Konturüberarbeitungen hervorgerufen.

Zusammenfassend sind viele der gebäudetypischen rechten Winkel gut erfasst. Das Zusammenspiel von Konturoptimierung und Beschränkung der Vertexverschiebungen durch Nebenbedingungen führt also meist zu den gewünschten Resultaten. Typische Problemfälle sind Gebäudekanten, deren Gradientenstärke nur schwach ausgeprägt ist. Eine Adaption an die falschen Bildkanten ist dann wahrscheinlich. Ferner können Fehler in vorhergehenden Arbeitsschritten bei der Nachbearbeitung nur bedingt ausgeglichen werden. Ist das Ausgangspolygon z. B. durch eine unzureichende Gebäudesegmentierung zu weit von der eigentlichen Gebäudekante entfernt, wird eine Adaption daran nicht gelingen. Führt man die Konturüberarbeitung auf benachbarten Gebäudeeinheiten aus, so kann es zu Überlappungen kommen. Stammen die Gebäudeeinheiten aus einer vorhergehenden Unterteilung,



**Abbildung 6.16:** Nachbearbeitung der automatisiert ermittelten Umrisspolygone (grün). Auf die nummerierten Fälle wird im Text eingegangen.

sind die Überlappungen allerdings meist gering.

### 6.7.2 Genauigkeit

Die Bewertung der nachverarbeiteten Gebäudekonturen im Hinblick auf die damit erzielbare Genauigkeit erfolgt, wie in den zwei vorangegangenen Kapiteln, anhand der in Gleichung 4.7 definierten Detektions-, Fehlrückweisungs- und Fehlalarmraten. Dazu wird eine automatisiert ermittelte Gebäudegrundfläche einer manuell vorgegebenen Referenz gegenübergestellt. Analog zur vorangegangenen qualitativen Beurteilung wird die Genauigkeit der Konturüberarbeitung anhand der künstlich verschlechterten wie auch der automatisch ermittelten Umrisspolygone berechnet.

#### Der künstlich erzeugte Datensatz

Die folgenden Betrachtungen beziehen sich auf die Nachbearbeitung der künstlich verschlechterten Ausgangspolygone des letzten Abschnitts. Berechnet man die Genauigkeit der Konturüberarbeitung auf Basis der (nicht-unterteilten) *Gebäudekomplexe*, so resultieren die im oberen Teil der Tabelle 6.1 angegebenen Werte. Die Angaben werden über alle Gebäudepixel ermittelt. Die zufällige Verände-

nung bewirkt in Bezug zur unveränderten manuellen Konturvorgabe eine Detektionsrate von 93,6 %. Eine anschließende Adaption (lokal) des Umrisspolygons an den lokalen Bildgradienten führt zu einer etwa einprozentigen Verbesserung der Genauigkeitsangaben. Da die Konturen nach der Überarbeitung geradliniger sind, scheinen sie die realen Gebäudeumrisse besser zu erfassen. Die zufälligen Verschiebungen erscheinen wesentlich ungeordneter; bei der Ermittlung der Genauigkeiten ist im Mittel jedoch kein großer Unterschied zu den überarbeiteten Konturen auszumachen.

Vorgehen	DR (%)	FRR (%)	FAR (%)
Konturvorgaben → Veränderung	93,59	6,41	7,92
Konturvorgaben → Veränderung → Konturüberarbeitung (K.Ü.) (lokal)	94,52	5,48	6,68
Konturvorgaben → Unterteilung → Veränderung	88,79	11,21	10,47
Konturvorgaben → Unterteilung → Veränderung → K.Ü. (lokal)	90,36	9,64	8,40
Konturvorgaben → Unterteilung → Veränderung → K.Ü. (global)	88,22	11,78	8,92
Konturvorgaben → Unterteilung → Veränderung → K.Ü. (hybrid)	89,85	10,15	8,51

**Tabelle 6.1:** Genauigkeitsangaben für die beschränkte Adaption von künstlich verschlechterten Gebäudeumrissen an Bildkanten. Die beiden ersten Zeilen geben die Genauigkeiten für Gebäudekomplexe an. Für die Genauigkeitsangaben der folgenden Zeilen werden die Gebäudekomplexe in einzelne Häuser unterteilt.

Auf eine nachfolgende Unterteilung der Gebäudesegmente in kleinere Einheiten durch das im letzten Kapitel vorgestellte Vorgehen wird verzichtet. Bei der hier gewählten Auswertung anhand aller Gebäudepixel führt eine Unterteilung zu keinen Veränderungen in Bezug auf die Genauigkeit, da sich die Menge der betrachteten Pixel nicht (wesentlich) verändert. Die Genauigkeiten müssten separat für jedes der unterteilten Gebäudesegmente ermittelt werden. Doch Unterteilungslinien für die überarbeiteten Gebäudepolygone können nicht auf die (zufällig verschobenen) Ausgangspolygone oder auf die Soll-Vorgabe angewendet werden. Deshalb steht keine geeignete Referenz zur Auswertung zur Verfügung. Eine Unterteilung der Soll-Vorgabe oder der daraus abgeleiteten Ausgangspolygone führt zu anderen Gebäudesegmenten. Für eine segmentweise Auswertung müssen jedoch die gleichen Segmente (in möglicherweise unterschiedlicher Ausformung) in allen Datensätzen vorhanden sein.

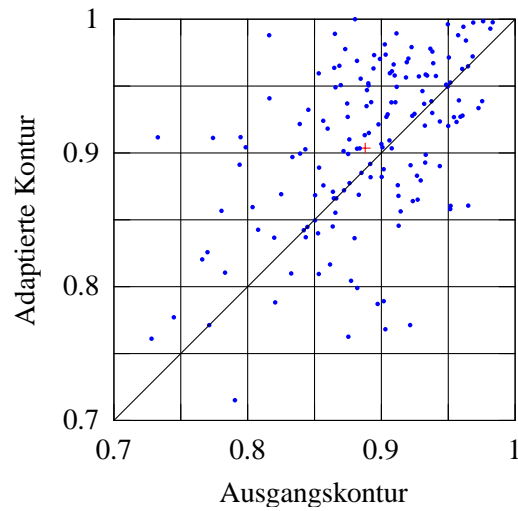
Die Auswertung der *hausweise* nachbearbeiteten Konturpolygone erfolgt etwas anders als die der Gebäudekomplexe. Für jedes der 161 im Testdatensatz vorkommenden Hauspolygone werden die Genauigkeitsangaben bestimmt und anschließend gemittelt. Daraus resultieren die im unteren Teil von Tabelle 6.1 wiedergegebenen Werte. Die Angaben sind mit den Genauigkeiten der Nachbearbeitung bei Gebäudekomplexen nicht direkt vergleichbar. Gebäudekomplexe haben eine wesentlich größere Ausdehnung als einzelne Häuser. Der von der Konturüberarbeitung betroffene Randbereich hat bei einzelnen Häusern also einen höheren Anteil an der Grundfläche und somit ein größeres

Gewicht bei der Ermittlung der Genauigkeit. Niedrigere Detektionsraten gegenüber vergleichbaren Angaben bei Betrachtung der Gebäudekomplexe sind die Folge. So sinkt die mittlere Detektionsrate bei den Ausgangspolygonen, also den zufällig verschobenen Konturvorgaben, auf 88,8 %. Dieser Wert wird nun als Referenz für die nachfolgende Konturnachbearbeitung herangezogen.

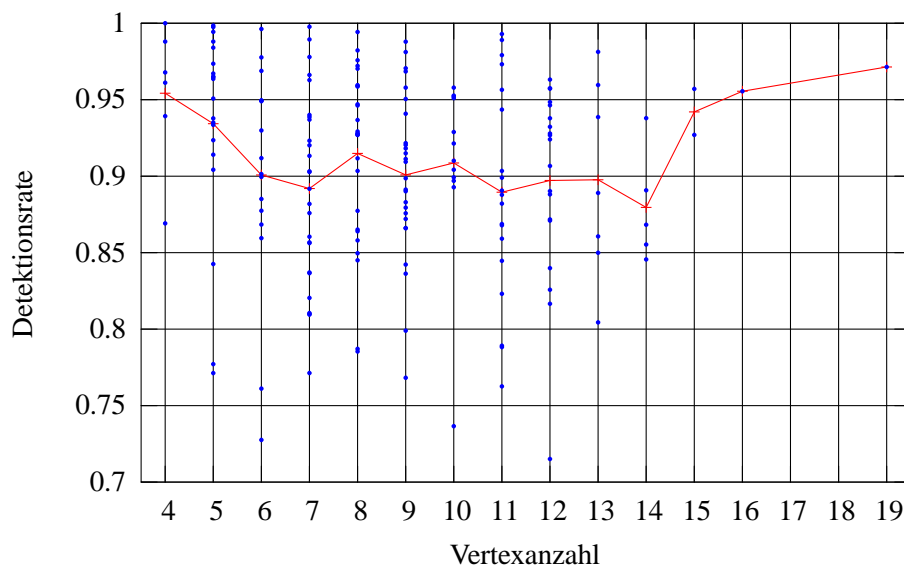
Um den Einfluss der Optimierungsstrategie zur Polygonadaption auf die erzielbare Genauigkeit zu untersuchen, werden bei diesem Datensatz auch unterschiedliche Optimierungsstrategien betrachtet. Daher resultiert in Tabelle 6.1 jeweils eine Zeile für die Genauigkeiten bei lokaler und globaler (Dynamische Programmierung) Optimierung sowie bei einer Mischform daraus. Bei letzterem Vorgehen werden die Vertexpositionen in der ersten Optimierungsschleife (Algorithmus 7) lokal optimiert, in der zweiten Schleife (Algorithmus 8) mittels Dynamischer Programmierung global. Nur die lokale Optimierungsstrategie führt mit einer Detektionsrate von 90,4 % zu einer Verbesserung der Genauigkeit gegenüber den Ausgangspolygonen. Die beiden anderen Vorgehensweisen führen in etwa zu den gleichen Genauigkeiten wie die nicht-überarbeiteten Ausgangskonturen. Der Einsatz der deutlich zeitaufwändigeren globalen Optimierungsmethode ist aus dieser Sicht also nicht gerechtfertigt. Für die Konturadaption mit Dynamischer Programmierung ist möglicherweise die Polygonkomplexität zu hoch. Gerade in der ersten Optimierungsschleife wird die Ausrichtung sehr vieler kurzer Polygonsegmente optimiert. Bei schlechten Ausgangskonturen ist dieser Optimierungsaufwand nicht nur unnötig, sondern evtl. auch zu einengend. Bei einer lokalen Optimierung sind Sprünge möglich, die nur lokal Vorteile bringen. Innerhalb eines global optimierten Adaptionsschritts können gut bewertete Pfade durch vorhergehende Vertices diese lokal gute Entscheidung jedoch verhindern. Eine globale Optimierung wäre somit zu starr für die in der Anfangsphase meist noch „unausgereifte“ Konstellation der Polygonstützpunkte. In dieser Phase scheint mehr Dynamik nötig, um sich aus Fehlkonstellationen zu lösen. Die geringfügige Verbesserung der mittleren Detektionsrate bei Verwendung des hybriden Ansatzes unterstützt diese Vermutung.

Die quantitative Veränderung der Detektionsraten bei Einsatz der Konturnachbearbeitung (unter Nutzung der lokalen Optimierungsmethode) ist in Abbildung 6.17 aufgetragen. Die Punkte stellen jeweils die Detektionsraten vor und nach der Überarbeitung dar. Befinden sich die Punkte über der Winkelhalbierenden hat sich die Detektionsrate verbessert. Eine Abhängigkeit der Verbesserung von der Ausgangsgüte ist nicht zu erkennen. Sowohl Verbesserungen als auch Verschlechterungen bezüglich der Detektionsrate sind auszumachen. Für den Großteil der Hauspolygone jedoch hat sich die Genauigkeit erhöht.

Um den Einfluss der Komplexität auf das Ergebnis der Konturüberarbeitung zu beleuchten, werden die bei lokaler Optimierungsstrategie hausweise ermittelten Detektionsraten in Abbildung 6.18 in Abhängigkeit von der Polygonkomplexität dargestellt. Als Maß der Komplexität wird die Anzahl der Stützpunkte des jeweils zugehörigen Soll-Polygons herangezogen. Die Detektionsraten aller Hauspolygone mit gleicher Vertexanzahl sind in der Darstellung entlang einer Senkrechten aufgetragen. Um daraus statistisch relevante Aussagen abzuleiten, ist der Stichprobenumfang zu gering. Jedoch lässt sich ein Trend erkennen: Je geringer die Komplexität, desto robuster funktioniert die Kontur-



**Abbildung 6.17:** Gegenüberstellung der hausweise berechneten Detektionsraten vor und nach der Konturüberarbeitung (blaue Punkte). Das rote Kreuz markiert die mittlere Detektionsrate.



**Abbildung 6.18:** Detektionsraten (blaue Punkte) in Abhängigkeit von der Komplexität der Konturlinie. Die mittlere Detektionsrate aller Häuser mit der gleichen Anzahl an Vertices ist durch die mit einer Linie verbundenen roten Kreuze dargestellt.

überarbeitung. Dadurch lassen sich bei einfachen Hauspolygonen höhere Genauigkeiten erzielen. Von den komplexen Polygonen (ab 15 Vertices) sind jeweils nur noch maximal zwei Repräsentanten vertreten — zu wenige, um aussagekräftig zu sein.

### Der Datensatz aus vorhergehenden Arbeitsschritten

Die folgenden Betrachtungen beziehen sich auf die Nachbearbeitung der automatisch ermittelten Ausgangspolygone, die sich aus vorhergehenden Modulen dieser Arbeit ergeben. Als Soll-Referenz wird bei der Betrachtung von Gebäudekomplexen dieselbe Handsegmentierung herangezogen wie bei der vorangegangenen Auswertung. Anders als in der letzten Auswertung wird bei den durch die Unterteilung entstandenen kleineren Gebäudeeinheiten ebenso diese Handsegmentierung verwendet. Die unterteilten Gebäude werden also hier *nicht* Einheit für Einheit bewertet, sondern die gesamte Grundfläche wird der einer Handsegmentierung gegenüber gestellt. Grund dafür ist, dass keine adäquate Soll-Vorgabe für die kleineren Gebäudeeinheiten ermittelt werden kann. Die Umrisspolygone spiegeln nicht immer einzelne reale Häuser wider, somit lässt sich eine solche Vorgabe nur schwer definieren.

Da mit einer globalen Optimierungsstrategie in der vorhergehenden Auswertung keine Verbesserungen erzielt werden, wird nachfolgend nur die lokale Optimierungsmethode gewählt.

Im Vergleich zu den Ausgangspolygonen bewirkt die Nachbearbeitung der Konturen der Gebäudeblocks keine Erhöhung der Detektionsrate. Wie dem oberen Teil von Tabelle 6.2 zu entnehmen ist, führen sowohl die ursprünglichen wie auch die überarbeiteten Gebäudepolygone zu einer Detektionsrate von 91,5 %. Die Angaben zur Genauigkeit der Ausgangspolygone entsprechen den Werten in Tabelle 4.3 in Abschnitt 4.4.3. Dort wird die Qualität der Gebäudesegmentierung ausgewertet.

Vorgehen	DR (%)	FRR (%)	FAR (%)
Konturlinien	91,51	8,49	8,68
Konturlinien → Konturüberarbeitung	91,46	8,54	6,59
Konturlinien → Unterteilung	89,84	10,12	8,68
Konturlinien → Unterteilung → Konturüberarbeitung	89,54	10,46	5,95

**Tabelle 6.2:** Genauigkeitsangaben für die beschränkte Konturoptimierung bei Gebäudekomplexen, die aus vorhergehenden Arbeitsschritten resultieren. Die beiden oberen Zeilen geben die Genauigkeiten für Gebäudekomplexe an, die beiden unteren Zeilen die Genauigkeiten für kleinere Gebäudeeinheiten.

Auch bei der Überarbeitung der kleineren Gebäudeeinheiten tritt keine Verbesserung hinsichtlich der Genauigkeit ein (s. unterer Teil von Tabelle 6.2). Die Messung erfolgt pixelweise unter Betrachtung der gesamten Gebäudeflächen. Die niedrigeren Detektionsraten im Vergleich zur vorangegangenen Messung bei Gebäudekomplexen liegt allein in der Umsetzung der Gebäudeunterteilungen begründet. Die Gebäudesegmente werden an den Trennlinien (Resultate nach Kapitel 5) mit einer ein Pixel breiten Linie in kleinere Segmente aufgeteilt. Die Ausgangspolygone ergeben sich dann

durch Konturieren dieser kleineren Segmente. Da die Soll-Vorgabe jedoch nicht unterteilt ist<sup>3</sup>, sinkt dadurch nur die Ist-Fläche und damit auch die Detektionsrate. Dieser Weg ist nur im Rahmen dieser Auswertung akzeptabel. Für eine routinemäßige Unterteilung der Gebäudesegmente bietet es sich an, die Konturpolygone der kleineren Gebäudeeinheiten durch einen Schnitt der Trennlinien mit dem Gebäudepolygon zu bestimmen.

Zusammenfassend ergibt sich folgendes Bild: Die Nachbearbeitung der Konturlinien führt meist zu einer visuell besser erscheinenden Beschreibung des Umrisspolygons von Gebäuden. Ursache sind eine meist geradlinigere und rechtwinkliger Umrisslinie wie auch die auf ein Minimum reduzierte Vertexanzahl. In der Detektionsrate spiegelt sich diese Verbesserung nicht bzw. nur bei ungeordneteren Ausgangspolygonen, wie sie durch zufälliges Vertexverschieben entstehen, wider. Ob die Gebäudekomplexe zuerst in kleinere Einheiten unterteilt werden und dann anschließend deren Kontur überarbeitet wird oder umgekehrt, ist bei Betrachtung der damit erzielbaren Genauigkeiten gleichwertig. Führt man die Konturüberarbeitung bei nicht-unterteilten Gebäudekomplexen durch, ist das Zustandekommen langer durchgängiger Umrisslinien, die das gesamte Gebäude beispielsweise von der angrenzenden Straße trennen, wahrscheinlicher.

### 6.7.3 Rechenzeit

Für die in diesem Kapitel vorgestellten Algorithmen werden nun die Resultate von Laufzeitmessungen auf dem in Abschnitt 1.5.1 angegebenen Rechensystem wiedergegeben. Grundlage der Konturüberarbeitung ist der in Abschnitt 6.2.2 behandelte Vorverarbeitungsschritt. Aus dem Bildgradienten wird durch Integration der initialen Gebäudekontur die lokale Gradientenstärke berechnet. Dieser Schritt braucht nicht vor der Nachbearbeitung jeder einzelnen Gebäudekontur durchgeführt werden, sondern kann vorab für die gesamte Bildszene erfolgen.

In Tabelle 6.3 sind sowohl die Laufzeiten dieses Vorverarbeitungsschritts als auch die zugehörige Komplexität der verwendeten Gebäudepolygone aufgeführt. Es wird unterschieden, ob die initialen Polygone von Gebäudekomplexen oder aber von einzelnen Häusern stammen. Die größeren Gebäudekomplexe umfassen im Allgemeinen mehr Stützpunkte als einzelne Häuser<sup>4</sup>. Dafür sind weniger Konturpolygone abzuarbeiten als bei der Betrachtung einzelner Häuser. Für die Bildszene mit  $2000 \times 2000$  Pixel liegt der Zeitaufwand zur Berechnung der lokalen Gradientenstärke im Bereich von 12–14 Minuten. Bei den 95 Konturpolygonen der Gebäudekomplexe läuft die Berechnung der lokalen Gradientenstärke um etwa eine Minute schneller ab als bei den 217 Polygonen der Einzelhäuser. Als Richtwert kann bei der genannten Bildgröße von einer mittleren Rechenzeit von etwa einer halben Sekunde pro Vertex ausgegangen werden. Zu dieser nicht unerheblichen Rechenzeit tragen

---

<sup>3</sup>Die Trennlinien für die automatisiert ermittelten Gebäudesegmente können aufgrund der anderen Segmentform auch nicht ohne Weiteres auf die vorgegebenen Gebäudesegmente übertragen werden.

<sup>4</sup>Die Unterscheidung in Polygone dieser beiden Kategorien bedeutet lediglich, dass die Gebäudekomplexe nicht weiter in kleinere Einheiten unterteilt werden. Sie umfassen jedoch genauso die Umrisse (beliebig kleiner) einzelner freistehender Häuser oder Innenhöfe.



vor allem die Umsetzung der Polygonstützpunkte in Linienpixel sowie die nachfolgenden Faltungen bei. Da die Berechnungen keine manuelle Interaktion erfordern, kann deren Ausführung zumindest zu beliebigen Zeitpunkten stattfinden.

Gebäudepolygone Typ	Anzahl	Vertexanzahl		Rechenzeit (s)		
		Mittel (Polygone)	Summe	Mittel (Vertices)	Mittel (Polygone)	Summe
Komplexe	95	11,14	1058	0,726	8,085	768,1
Einzelhäuser	217	8,62	1870	0,440	3,790	822,4

**Tabelle 6.3:** Laufzeitangaben für die Berechnung der lokalen Gradientenstärke. Bei den angegebenen Mittelwerten wird über die jeweils in der Kopfzeile aufgeführten Größen gemittelt.

Zur Bewertung des Laufzeitverhaltens der Konturüberarbeitung wird der bereits in Abschnitt 6.6 zur Visualisierung des Programmablaufs verwendete Datensatz mit einem zusammengesetzten Gebäude verwendet. Die Untersuchung erfolgt also nur an *einem* Gebäudepolygon. Für größere Datensätze kann die Laufzeit dann anhand der Gebäudeanzahl bzw. der mittleren Anzahl an Vertices abgeschätzt werden.

Die Pixelauflösung der Bilddaten des Testdatensatzes beträgt 50 cm. Das initiale Polygon weist nach Vereinfachung 21 Stützpunkte auf; das zugehörige Gebäude hat eine Größe von etwa 780 m<sup>2</sup>.

Tabelle 6.4 gibt die Ergebnisse der Zeitmessung wieder. Für jeden Arbeitsschritt wird die Rechenzeit wie auch die Anzahl der Vertices, aus denen das Gebäudepolygon in diesem Schritt besteht, angegeben. Für die beiden Schritte zur Forcierung der Nebenbedingungen wird zusätzlich die Anzahl der Constraints angegeben. Anzumerken ist, dass die Anzahl der Vertices und der Constraints sich auf zweidimensionale Punkte bezieht. Für die Zustandsvektor-Notation gemäß Gleichung 6.21 sind diese Angaben zu verdoppeln.

Arbeitsschritt	Rechenzeit (ms)			Vertices [/ Constraints] (#)		
	Min.	Mittel	Max.	Min.	Mittel	Max.
Rechtw.-Constraint	1,2	4,9	19,9	26 / 1	26 / 4,3	26 / 6
Rechtw.- und Ori.-Constraint	1,8	1,9	2,0	10 / 2	10 / 5	10 / 6
Polygonadaption (lokal)	849,7	1731,0	2407,9	10	21,3	30
Polygonadaption (global)	16510,7	33689,5	48891,5	10	21,3	30
Beschränkung bzgl. Constraint	0,3	0,4	0,4	10	20,7	26
Dynam. Vertexverwaltung	20,8	73,9	191,1	10	10,4	14

**Tabelle 6.4:** Laufzeitangaben für die beschränkte Konturoptimierung. Aus allen Iterationen des Gesamtdurchlaufs werden die minimalen/mittleren/maximalen Werte pro Iteration ermittelt.

Die in Abschnitt 6.5 erläuterten Schritte zur dynamischen Verwaltung der Vertices sind hier zu einem Schritt zusammengefasst. Dabei ist zu erwähnen, dass nicht alle dieser Schritte in jeder Iteration zum Einsatz kommen. So wird das Ausprägen rechter Winkel nur im (relativ seltenen) Bedarfsfall angewendet.

Im Vergleich zur eigentlichen Polygonadaption fällt die Laufzeit der anderen Schritte nicht wesentlich ins Gewicht. Insbesondere die Umsetzung der Nebenbedingungen ist sehr schnell zu verwirklichen ( $< 20$  ms). Grund dafür ist, dass die gesamten Rechenoperationen nur auf den Polygonkoordinaten basieren. Im Gegensatz zur Polygonadaption wird dabei nicht auf den lokalen Bildgradienten zurückgegriffen.

Die Adaption an die Bildkanten ist vor allem deshalb so zeitintensiv, da die Polygongüte anhand *aller Pixel entlang der Polygonlinie* ermittelt wird. Nur so wird sichergestellt, dass die Konturlinie nicht nur punktuell bewertet wird. Bei der Implementierung Aktiver Konturen ist es weitaus verbreiteter, die externe Energie nur an den Polygonstützstellen auszuwerten. Als Kompromiss wäre eine punktuelle Auswertung mit einer erhöhten Vertexdichte denkbar.

Vergleicht man bei der Polygonadaption die lokale mit der globalen Optimierungsmethode durch Dynamische Programmierung, so wird der zu Beginn dieses Abschnitts erwähnte zeitliche Unterschied deutlich. Der globale Optimierungsansatz führt im Mittel zu einer fast 20-fach längeren Laufzeit (bei einer mittleren Anzahl von 21,3 Polygonstützpunkten). Dieser zeitliche Mehraufwand ist gemäß der Genauigkeitsbewertung im letzten Abschnitt nicht gerechtfertigt.

## 7 Zusammenfassung und Ausblick

In diesem Abschlusskapitel werden die wesentlichen Ergebnisse der Arbeit zusammengetragen (Abschnitt 7.1). In Form eines Ausblicks (Abschnitt 7.2) zu den methodischen Teilen dieser Arbeit (Kapitel 3 bis 6) werden mögliche Anknüpfungspunkte für weitere Untersuchungen vorgestellt.

### 7.1 Zusammenfassung der Arbeit

Im Rahmen dieser Arbeit wurde ein Verarbeitungsschema entwickelt, um von hochauflösenden Bilddaten zu der Konturlinie von Objekten sowie zugeordneten Attributen zu gelangen. Dabei steht die Extraktion von Gebäuden im Vordergrund. Diese Arbeit stellt somit einen wichtigen Beitrag für die Automatisierung des Übergangs von Pixelinformation zu interpretierter abstrakter Information dar. Die zunehmende Verbreitung geografischer Informationssysteme und die enormen Datenmengen hochauflösender Bilddaten erfordern derartige Automatisierungsschritte in der Bildinterpretation.

Nach einer Hinführung zur Thematik und den Ausgangsdaten der Arbeit wird in Kapitel 1 ein Literaturvergleich angestellt. Aus diesem ergeben sich folgende wesentliche Randbedingungen für die Entwicklung des Vorgehens zur automatisierten Gebäudeextraktion:

- Nutzung der hohen Bildauflösung und der komplementären Information aus DOM und Multispektraldaten
- Prinzipielle Erweiterbarkeit auf andere Objektkategorien als Gebäude
- Möglichst wenige Annahmen über die Gebäudebeschaffenheit (Form und Material, also spektrale Charakteristik)
- Anwendbarkeit insbesondere im Stadtbereich mit Häuserblocks

Die hohe geometrische Auflösung der Ausgangsdaten wird zur Extraktion detaillierter Strukturen (Brandmauern) und zur genauen Erfassung der Konturlinien der Gebäudedächer benötigt. Aus der komplementären Information von DOM und Multispektraldaten werden Objektattribute abgeleitet, die aus dem jeweils anderen Datenprodukt nicht oder nur schwer und ungenau zu bestimmen wären. Das DOM wird zur Bestimmung der Objekthöhen herangezogen, die Multispektraldaten zur Identifikation von Vegetation, die gute Kantenqualität der Orthobilder zur Korrektur der Gebäudeumrisse.

Die Erweiterbarkeit der Gebäudeextraktion auf eine verallgemeinerte Objektextraktion wird über eine Modularisierung des Extraktionsvorgangs erreicht. In vier aufeinander folgenden Schritten wird

der Objektbegriff durch Hinzuziehen weiterer Attribute jeweils verfeinert. Die Extraktion anderer kartografisch relevanter Objekte wird durch Hinzufügen weiterer Module ermöglicht. In einer hierarchisch strukturierten Objektextraktion wird der Implementierungsaufwand durch Nutzung der Segmentierungsergebnisse der bereits vorliegenden Module minimiert.

Die Gebäudesegmente werden durch eine vorgeschaltete Höhen-, Vegetations- und Schattensegmentierung nur indirekt bestimmt. Da eine explizite Formulierung von Vorgaben für die Form, die spektrale oder textuelle Zusammensetzung der Gebäudesegmente somit umgangen wird, sind die Gebäude diesbezüglich auch nur minimal beschränkt. Ähnlich verhält es sich bei der späteren Überarbeitung der Gebäudeumrisse. Dort wird die Gebäudegeometrie nur dann beschränkt, wenn die umzusetzende Konstellation der Konturlinie bereits annähernd in den ursprünglichen Gebäudesegmenten erkennbar ist. Im Unterschied zu modellbasierten Ansätzen wird also keine Gebäudegeometrie vorgegeben, sondern in geeigneten Fällen eine gebäudetypische Geometrie forciert.

Gebäudekomplexe werden nicht als Ganzes extrahiert, sondern aufgeteilt in einzelne Häuser, denen Besitzverhältnisse zugeschrieben werden können. Dadurch ist der vorgestellte Ansatz nicht beschränkt auf einzelne freistehende Häuser oder die Extraktion von Häuserblocks. Die Verwendung typischer Stadtscenen mit mehreren Gebäudekomplexen als Testdatensätze unterstreicht dies.

Der in Kapitel 2 vorgestellte Erdfernerkundungssensor HRSC-A liefert die geometrisch hochaufgelösten Ausgangsdaten dieser Arbeit. Die hohe geometrische Qualität der Daten machen sie zur idealen Grundlage für Interpretationsaufgaben mit hohen Genauigkeitsanforderungen wie Planungsarbeiten oder die Erzeugung, Ergänzung bzw. Aktualisierung amtlicher Geobasisdaten. Die geometrische Genauigkeit wird erzielt durch die genaue Anbindung jeder aufgenommenen Bildzeile an die Position und Lage während der Befliegung, was eine Fluglagekorrektur der Multispektraldaten ermöglicht. Die stereophotogrammetrische Zuordnung der fünf panchromatischen Kanäle führt zu einem DOM. Dies wiederum ist die Grundlage für eine Orthokorrektur der Bilddaten. Die daraus resultierenden multispektralen Orthobilder sowie das DOM bilden die Datenbasis dieser Arbeit. Zusatzinformation, wie Gebäudegrundpläne oder bereits vorliegende Klassifikationsergebnisse, werden nicht verwendet.

Für die Methodik zur Gebäudeextraktion werden die hohe geometrische Auflösung, die Multispektral- und die Orthoeigenschaften der Bilddaten sowie die Höheninformation aus dem DOM genutzt. Basis für die NDVI-Berechnung in Kapitel 4 sind zwei Kanäle, die den Wellenlängenbereich um 700 nm flankieren. Daten aus anderen Quellen mit diesen Eigenschaften können somit ebenso zum Einsatz kommen.

Die Bestimmung der Objekthöhen erfolgt durch eine Normierung des DOM auf Grundhöhe. Grundlage dafür ist die Berechnung eines Terrainmodells, wozu in Kapitel 3 zwei Ansätze vorgestellt werden. Bei einem der Ansätze erfolgt die Ermittlung der erhöhten Regionen durch die benutzerbestimmte Wahl von Abstands- und Höhenschwellen. Im Unterschied zu morphologischen Ansätzen bleiben die Höhenangaben aus dem DOM bei den als Gelände eingestuftten Regionen unmodifiziert erhalten. Lediglich die Grundhöhen der erhöhten Regionen werden anhand der Höhen

umgebender Grundregionen interpoliert. Im zweiten Ansatz steht ein höherer Automatisierungsgrad im Vordergrund. Erhöhte Regionen werden durch schrittweises Verkleinern einer sie umgebenden Referenzregion in Form einer Auflösungshierarchie ermittelt. Steigt innerhalb dieser Referenzregion die Grundhöhe im Vergleich zur vorhergehenden Iteration sprunghaft an, wird auf das Vorliegen erhöhter Objekte geschlossen. Sie bekommen die zuletzt bestimmte Grundhöhe zugewiesen. Hierbei wird die Grundhöhe also nicht interpoliert, sondern sie stammt direkt von umgebenden Grundbereichen ab. Das erste Verfahren lässt sich durch Inspektion der Zwischenprodukte und eine entsprechende Feinkorrektur der Parameterwahl sehr genau an das vorliegende Terrain anpassen. Die an ausgewählten Punkten in einem Testdatensatz erzielten Genauigkeiten liegen bei 40 bzw. 110 cm für das erste bzw. zweite Verfahren.

Aus dem so bestimmten Terrainmodell ergeben sich durch Subtraktion vom DOM unmittelbar Objekthöhen. Eine Hörschwelle führt zu erhöhten Regionen. In Kapitel 4 wird eine Methode zur Differenzierung dieser erhöhten Regionen in Bäume und Gebäude vorgestellt und diskutiert. Durch ein NDVI-basiertes Verfahren wird zunächst die Vegetationszugehörigkeit bewertet. Insbesondere bei Nutzung von Daten der HRSC-A wird aufgrund der für eine Vegetationssegmentierung ungünstigen Konstellation des Rotkanals die Integration weiterer Kanäle vorgeschlagen. Anhand einer Farbtonsegmentierung der bisher ungenutzten Multispektralkanäle werden vorläufige Baumsegmente bestimmt. Das NDVI-Kriterium entscheidet über die Akzeptanz des Segments. Von den nicht als Vegetation eingeordneten Segmenten werden die durch eine einfache Schattendetektion ermittelten Schattenregionen am Rand der Segmente ausgeschlossen. Die verbleibenden Regionen gelten nach einer morphologischen Nachverarbeitung als Gebäudesegmente. Die in der Erdfernerkundung verbreitete Vegetationssegmentierung kann durch das vorgestellte Vorgehen auch auf hochauflösende Bilddaten, insbesondere auf Daten der HRSC übertragen werden. Damit können Bäume mit über 80 %, Gebäude mit über 90 % Genauigkeit erkannt werden.

In Kapitel 5 wird die Aufteilung der Gebäudesegmente, die durchaus auch größere Wohnblocks darstellen können, in einzelne Häuser beschrieben. Dazu wurde ein neuartiger, auf der für Brandmauern typischen Gradienteninformation (sowohl die Stärke als auch die Orientierung des Bildgradienten sind ausschlaggebend) beruhender Ansatz ausgearbeitet. Die Skeletpixel eines Gebäudesegments stellen die Ausgangspunkte für Hypothesen zur Unterteilung anhand von Trennlinien dar. Die Orientierung benachbarter Segmentrandpixel bestimmt die Ausrichtung der Trennlinie. Durch das Überprüfen geometrischer wie auch radiometrischer Kriterien werden nun schrittweise Hypothesen ausgeschlossen. Anhand eines Qualitätskriteriums wird eine Rangordnung der Hypothesen eingeführt. Entsprechend dieser Ordnung werden die Trennlinien unter Beachtung von Konsistenzkriterien umgesetzt, also das Gebäudesegment unterteilt. Auf diese Art wird an einem Testdatensatz eine Übereinstimmung von etwa 60 % mit manuell vorgegebenen Linien erreicht. Werden die Gebäudeunterteilungen manuell überarbeitet, sind die automatisch detektierten Unterteilungslinien eine große Hilfe. Andernfalls ist die erreichbare Genauigkeit noch zu gering.

Schließlich wird in Kapitel 6 bei Betrachtung der Hauskontur die geometrische Qualität der Bild-

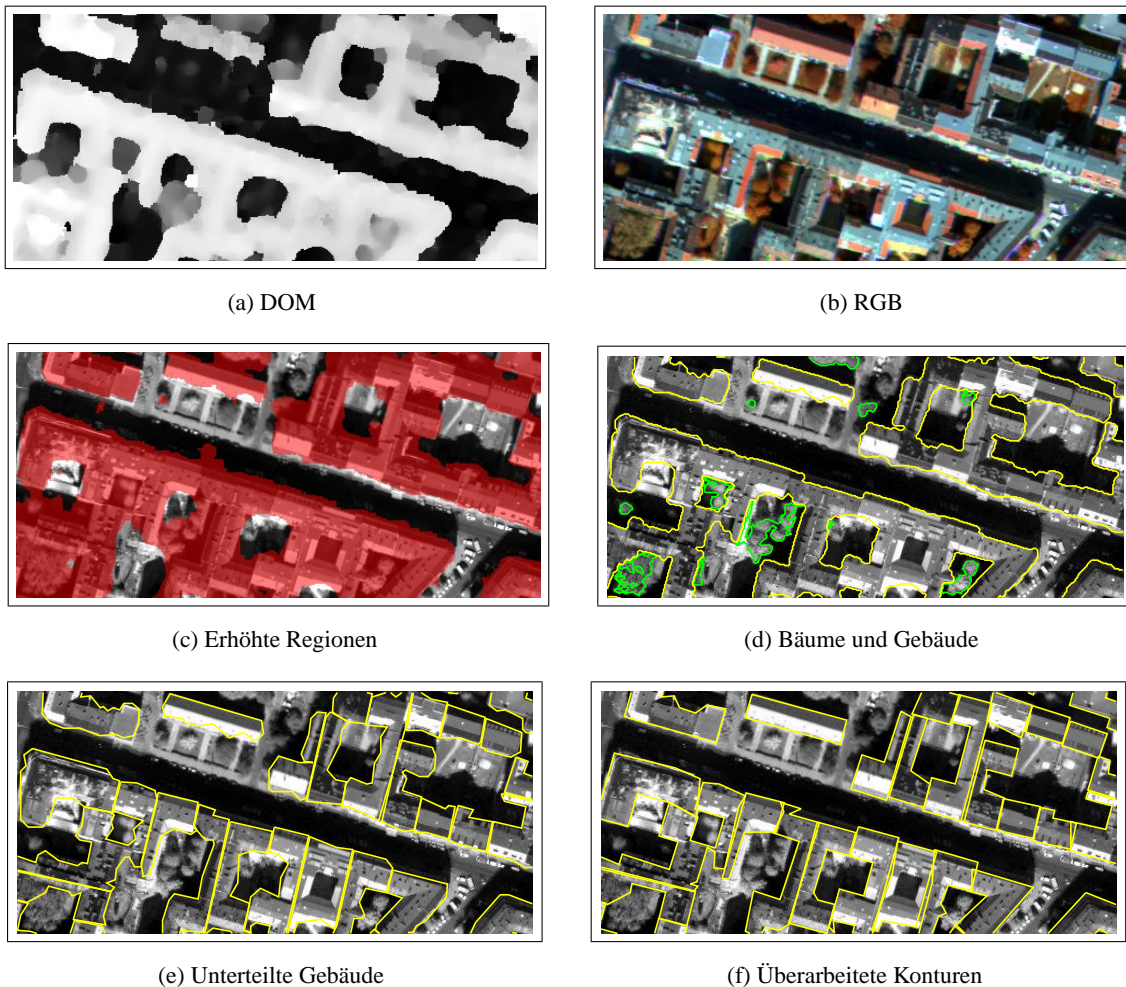
kanten in den Orthobildern genutzt. Durch einen Adaptionprozess wird die Form der Haussegmente auf Grundlage des Bildgradienten neu ausgerichtet. Gegenüber Ansätzen mit Aktiven Konturen zeichnet sich dieses Vorgehen durch folgende Punkte aus: Zur Bewertung der Konturqualität werden *alle* an der Konturlinie beteiligten Bildpunkte herangezogen, nicht nur die Vertices. Grundlage für die Adaption ist sowohl die Gradientenstärke als auch die Gradientenorientierung. Dadurch wird der Einfluss von Kantenpixeln, deren Orientierung nicht mit der der Konturlinie übereinstimmt, unterdrückt. Ferner wird auf die bei Aktiven Konturen üblichen Regularisierungskräfte innerhalb der Adaptionssphase verzichtet. Die Regularisierung des Umrisspolygons findet in Form gesonderter Schritte außerhalb der eigentlichen Optimierung statt. Dabei wird auch die Anzahl der Vertices variiert.

Zur Forcierung gebäudetypischer Geometrien wie Rechtwinkligkeit oder Parellelität werden die Adaptionsschritte anhand des von FUA UND BRECHBÜLER (1996) für Aktive Konturen eingeführten Vorgehens über Nebenbedingungen beschränkt. Dabei ist die Umsetzung der Nebenbedingungen ideal auf den iterativen Optimierungsablauf abgestimmt.

Für die Polygonadaption hat sich eine lokale Optimierungsstrategie in Bezug auf die damit verbundene Laufzeit, aber auch auf die erzielbare Genauigkeit bewährt. Die Detektionsrate bleibt dabei auf dem hohen Niveau der Gebäudesegmentierung. Die Umrisse der Häuser sind nun gebäudetypischer, also geradliniger und mit ausgeprägten rechten Winkeln. Diese Verbesserung der Konturqualität erleichtert die Übernahme in ein GIS. Bei Visualisierungsanwendungen wie 3-D-Darstellungen vermitteln die nun geradlinigeren Umrisslinien einen deutlich realistischeren Eindruck. Zudem ist die Zahl der Polygonstützpunkte auf ein Minimum reduziert. Damit ist der Aufwand bei einer manuellen Nachbearbeitung der Konturen drastisch reduziert.

Abbildung 7.1 veranschaulicht die wesentlichen Schritte dieser Arbeit an Bildbeispielen. Da in diesem Bildausschnitt ein ebenes Gelände in der Stadt erfasst ist, wird auf die DOM-Normierung verzichtet. Die erhöhten Regionen ergeben sich bereits aus dem DOM (Abb. 7.1(c)). Anhand der multispektralen Information werden Vegetationssegmente identifiziert. Schließt man die Vegetationssegmente und Schattenregionen am Segmentrand von den erhöhten Regionen aus, so gelangt man zu Gebäudesegmenten (s. Abb. 7.1(d)). Diese Gebäudesegmente werden nun anhand der Gradienteninformation eines Bildkanals unterteilt in kleinere Einheiten (Abb. 7.1(e)). Abschließend werden die Konturen der Häuser bezüglich der Übereinstimmung mit Bildkanten, deren Geradlinig- und Rechtwinkligkeit sowie der Anzahl der Polygonvertices optimiert (Abb. 7.1(f)).

Durch die aus den vorgestellten Methoden resultierenden objektspezifischen Attribute und die zugehörige Objektkontur ist zum einen der Weg für eine objektbasierte Interpretation geebnet. Zum anderen stehen damit bereits wichtige Eingangsdaten für ein Geoinformationssystem, insbesondere für den urbanen Bereich, zur Verfügung. Bestrebungen zur Automatisierung von Abläufen in der Interpretation von Bilddaten der Erdfernerkundung sind und werden verstärkt das Feld intensiver Forschung sein. Diese Arbeit stellt einen Beitrag dazu dar.



**Abbildung 7.1:** Bildbeispiel zum Ablauf der Objektextraktion. Dabei werden folgende Farbkodierungen verwendet. Erhöhte Regionen: rot, Bäume: grün, Gebäude: gelb.

## 7.2 Ausblick

Nachfolgend werden Möglichkeiten zur Weiterentwicklung der in den Kapiteln 3 bis 6 vorgestellten Extraktionsschritte aufgezeigt und diskutiert. Die Untergliederung in Abschnitte entspricht der kapitelweisen Trennung der jeweiligen Methoden dieser Arbeit. Vorab wird auf Aspekte eingegangen, die nicht methodenspezifisch sind.

### Vergleich mit bestehenden Datenbeständen

Die in dieser Arbeit vorgestellten Verfahren wurden anhand manuell ermittelter Vorgaben bewertet. Als Kriterium für den Einsatz hochauflösender digitaler Bilddaten in Kombination mit einer automatisierten Auswertung ist auch ein Vergleich mit bisherigen Datenbeständen notwendig. Eine reine Ge-

genüberstellung z. B. hausweise ermittelter Detektionsraten wird nicht ohne Weiteres möglich sein. Da die Datensätze in den meisten Fällen nicht zum gleichen Zeitpunkt oder mit dem gleichen Sensor erhoben werden, sind bereits unabhängig von der Auswertemethode andere Ausprägungen der Objekte zu erwarten. Zur Erörterung dieser Fragestellungen stellt KRÜGER (2002) einen Vergleich von Gebäudepolygonen aus der Automatisierten Liegenschaftskarte Berlin mit Bildkanten von HRSC-A Daten an. Neben der genannten Problematik der unterschiedlichen Aufnahmezeit und -sensorik wird auch auf Sonderfälle eingegangen, die einen direkten Vergleich erschweren. So können von Bäumen überdeckte Gebäude aus der Orthoperspektive nicht ausgemacht werden. Für Tiefgaragen, die in der Liegenschaftskarte ebenfalls als „Gebäude“ aufgeführt sind, gilt dasselbe.

In Abschnitt 1.1 wird einführend das Problem der Aktualisierung von Geodaten erwähnt. Die Konturlinien aus den neuen Datensätzen stimmen meist nicht völlig mit denen der bereits vorhandenen Datenbasis überein. Ein adaptives Vorgehen, wie es in Kapitel 6 beschrieben wurde, bietet sich für einen *automatischen Abgleich der Datensätze* an — mit den zuvor erwähnten Einschränkungen. Statt auf das DOM kann in einem solchen Fall auf den bereits vorhandenen Datenbestand als Grundlage für initiale Gebäudekonturen zurückgegriffen werden. Eine Konturüberarbeitung führt dann zu einer Angleichung an die genaueren Daten.

### **Ermittlung der Geländehöhe**

Die in Kapitel 3 vorgestellten Ansätze zur Ermittlung der Geländehöhe sind in erster Linie mit dem Ziel der Gebäudedetektion entworfen worden. Durch die Anwendung in urbanen Gebieten kann davon ausgegangen werden, dass über die gesamte Szene verstreut immer wieder Grundpunkte als Referenz vorhanden sind. Ferner erscheint eine lineare Interpolation der Grundhöhe hier adäquat. In bewaldeten Gebieten mit Baumkronendächern, die sich weiträumig erstrecken, müssen diese beiden Annahmen nicht gegeben sein. Zum einen erlauben in einem solchen Fall die — wenn überhaupt vorhandenen — meist kaminartig engen Blicke auf den Waldboden keine photogrammetrische Rekonstruktion der Bodenhöhe. Zum anderen kann eine lineare Interpolation über weit ausgedehnte Flächen die Geländemorphologie nicht erfassen.

### **Trennung von Vegetation und Gebäuden**

Da die Trennung von Vegetation und Gebäuden anhand spektraler Charakteristika erfolgt, ist auf die radiometrische Qualität der Bilddaten besonderen Wert zu legen. Die Kanäle der HRSC sind radiometrisch jedoch nur relativ zueinander kalibriert (s. Abschnitt 2.3.2). Absolute Messungen sind in den Bilddaten daher nicht möglich. Die Methoden in Kapitel 4 umgehen dies durch Bildung von Kanalaratios. Radiometrische *Normierungen anhand physikalischer Strahlungsmodelle* versprechen genauere Ergebnisse. Beispielsweise seien hier die Korrektur der atmosphärischen Streuung oder die photometrische Korrektur des je nach Geländeneigung blickwinkelbedingt unterschiedlichen Rückstrahlverhaltens genannt.



Wie sich bei der Auswertung in Abschnitt 4.4 gezeigt hat, ist die Schattenproblematik ein nicht zu vernachlässigendes Thema. Mehrere Bäume können nicht detektiert werden, da sie gänzlich durch Schatten abgedeckt sind. Auch beschattete Teile von Gebäuden werden, falls sie sich am Rande des Gebäudesegments befinden, oft nicht detektiert. Eine *aufwändigere Schattensegmentierung* könnte dazu beitragen, diese Probleme zu verringern. Beispielsweise kann dies durch Integration des DOM in die Schattendetektion erfolgen. In der Computergrafik wird das DOM dazu entsprechend dem Sonnenstand zur Aufnahmezeit (bei HRSC-Produkten durch GPS-Messung automatisch erfasst) künstlich beleuchtet. Für einen umfassenden Überblick über Schattierungsalgorithmen sei auf WOO ET AL. (1990) verwiesen. Generell gilt, dass Schatten durch Schattierungsverfahren nur in der Genauigkeit des DOM projiziert werden können. Diesbezüglich sind bei Verwendung eines HRSC-Höhenmodells die in Abschnitt 2.5.2 erläuterten Interpolationsungenauigkeiten zu berücksichtigen. So beziehen STEVENS ET AL. (1995) die DOM-Fehler durch einen probabilistischen Ansatz in einen Schattierungsalgorithmus ein.

Da die Ergebnisse der Schattendetektion zur Korrektur der Gebäudesegmente verwendet werden sollen, ist eine den realen Bilddaten möglichst gut entsprechende Schattensimulation erforderlich. Deshalb müssten die anhand einer Schattierung des DOM ermittelten Schattenregionen in einem weiteren Schritt mit den Bilddaten abgeglichen werden.

Oftmals liegen bei einer Auswertung eines Befliegungsgebietes *weitere Informationsquellen* vor. Sind entsprechende Karten oder bereits interpretiertes Bildmaterial anderer Befliegungen vorhanden, bietet sich deren Integration in den Unterscheidungsprozess Vegetation-Gebäude an. Ist beispielsweise ein Segment in einer Kartierung als Gebäude gekennzeichnet, so sollte die Wahrscheinlichkeit für eine Gebäudezuordnung im Segmentierungsprozess erhöht werden.

### Unterteilung von Gebäudekomplexen

Die in Kapitel 5 vorgestellten Schritte stellen einen ersten Ansatz zur Unterteilung von Gebäudesegmenten dar. Der Einfachheit halber werden dabei nicht alle real möglichen Unterteilungen zugelassen. *Trennlinien, die im Inneren des Gebäudesegments enden*, sind dann sinnvoll, wenn weitere Linien für eine vollständige Abtrennung existieren. Analog zu der schrittweisen Realisierung von Trennlinien in Innenhöfe (Abschnitt 5.7) kann dies überprüft werden. Da jedoch Definition 5.1 durch Linien, die im Segmentinneren enden, verletzt ist, müssen neue Wege zur Generierung der Trennlinienkandidaten gefunden werden.

Die zweite erwähnenswerte Einschränkung ist der *Ausschluss sich kreuzender Trennlinien*. Dazu ist anzumerken, dass der Ausschluss erst zum spätest möglichen Zeitpunkt (bei der Umsetzung der Trennlinien) erfolgt. Das Verwerfen sich kreuzender Linien findet dadurch unter Einbeziehung der größtmöglichen Informationsmenge statt. Lässt man sich kreuzende Trennlinien generell zu, so ist eine aufwändigere Verwaltung der Mindestsegmentgröße sowie die Möglichkeit zur nachträglichen Modifikation der kreuzenden Linien notwendig. Nach Verwirklichung einer Linie wird es dann

möglich, dass eine kreuzende Linie sich nur bis zu der gerade realisierten Trennlinie erstreckt. Dieser letzte Modifikationsschritt kann entfallen, wenn — wie im Absatz zuvor besprochen — Linien auch im Segmentinneren enden dürfen.

Bei der in Abschnitt 5.4 behandelten *Verkürzung der Linienenden* können aus Trennlinien Polygonzüge mit bis zu vier Stützpunkten werden. Die Verkürzung hängt ausschließlich von der Lage der Linie im Gebäudesegment ab. Die folgenden Schritte betrachten jedoch aus Vereinfachungsgründen nur den inneren Teil dieses Polygonzugs, also die eigentliche Linie. Eine Verbesserung der Genauigkeit ist zu erwarten, wenn zum einen weitere Aspekte, wie z. B. die Gradientenstärke und -orientierung, bei der Verkürzung der Linien integriert werden. Zum anderen, falls nachfolgend, z. B. bei der Bewertung der Linien, der komplette Polygonzug einbezogen wird.

Eine weitere Verbesserung kann durch die *Integration von Kontextwissen* erzielt werden. So liegt es nahe, die Symmetrie der großteils regelmäßig aufgebauten Gebäudekomplexe, einzubeziehen. Beispielsweise können lange gut bewertete Linien als Hauptachsen des Gebäudesegments betrachtet werden. Treten zwei Linien spiegelbildlich zu einer der Hauptachsen auf, so kann dies zur Aufwertung dieses Linienpaars beitragen.

Wird die automatische Unterteilung der Gebäudesegmente als *Vorstufe für eine manuelle Auswahl und Korrektur* genutzt, kann die Überprüfung nachbarschaftlicher Beziehungen der Linien reduziert werden. Der Anwender erhält so eine größere Auswahl an Trennlinien, aus der er sich für eine Untermenge entscheidet und diese ggf. ergänzt. Dies bedeutet auch, dass der Anwender die Instanz repräsentiert, die für die Konsistenz der Trennungen sorgt. In den automatisch generierten Trennlinien könnten so beispielsweise Kreuzungen durchaus existieren.

### **Nachbearbeitung der Gebäudekonturen**

Der Genauigkeitszugewinn durch die Anwendung der in Kapitel 6 beschriebenen Nachbearbeitung der Gebäudekonturen ist nur begrenzt (s. Abschnitt 6.7.2). Ansätze zur Qualitätsverbesserung haben in den zuvor ablaufenden Schritten eine größere Auswirkung. Dennoch gibt es auch hier zahlreiche Möglichkeiten zur Weiterentwicklung. Insbesondere das *Zusammenspiel mit der in Kapitel 5 behandelten Unterteilung der Gebäudesegmente* bietet Potenzial für Verbesserungen. So sollten Trennlinien, die einen Gebäudekomplex unterteilen, durch die Konturüberarbeitung nicht mehr oder nur noch minimal bewegt werden. Die Trennlinien sind per Definition bereits an den Bildgradienten angepasst. Die Fixierung des Polygonsegments, das die Trennlinie darstellt, könnte beispielsweise über eine entsprechende Nebenbedingung implementiert werden. Ein solches Vorgehen würde sehr gut mit der Forcierung gebäudetypischer Geometrien nach Abschnitt 6.4 harmonisieren. Durch eine Fixierung der Trennlinie während der Nachbearbeitung wäre auch ausgeschlossen, dass die betreffenden Polygonsegmente zweier benachbarter Gebäude unterschiedlich ausgerichtet werden und es so zu Überlappungen kommt.

Die Konturüberarbeitung wird für jeden Gebäudeumriss separat durchgeführt. Da Gebäude eines

Straßenzuges oftmals die gleiche Orientierung aufweisen, bietet sich diese *gebäudeübergreifende Hauptorientierung* zur Integration in den Optimierungsprozess an. Ebenso eignen sich die Konturlinien des Gebäudeumrisses und des zugehörigen Innenhofs aufgrund der meist gleichen Ausrichtung zu einer gekoppelten Nachbearbeitung.

Die Adaption der Polygonstützpunkte an den lokalen Bildgradienten ist mit einer gleich großen Nachbarschaft für alle Vertices implementiert. Da jedem Vertex eine Gütefunktion zugeschrieben ist, lässt sich die *Größe des Adaptionsbereichs an ein Qualitätsmaß koppeln*. Für gut bewertete Vertices genügt zur Feinadaption eine kleine Nachbarschaft; schlechter bewertete Stützpunkte hingegen haben einen höheren „Verbesserungsbedarf“.

Eine Variation des Adaptionsbereichs kann ebenfalls durch *Veränderung der Bildauflösung* erzielt werden. Beginnend bei einer geringen Auflösung wird eine grobe und schnelle Adaption an den Gebäudeumriss erreicht. Nach einer schrittweisen Erhöhung der Bildauflösung wird die Konturüberarbeitung immer feiner umgesetzt. GÜLCH (1990) und WHITTEN (1993) haben dazu die Scale-Space-Theorie von LINDBERG (1990) mit Aktiven Konturen verknüpft.



# Symbol- und Abkürzungsverzeichnis

$A$	Jakobi-Matrix der Nebenbedingungen, Seite 142
$A_{\text{CCD}}$	Fläche eines CCD-Elements, Seite 32
ALK	Automatisierte <i>LiegenschaftsKarte</i> , Amtliches Geoinformationssystem zur Verwaltung von Liegenschaften, Seite 2
BRDF	Bidirektionale Reflektanzverteilungsfunktion ( <i>'Bidirectional Reflectance Distribution Function'</i> , Seite 27
CCD	Elektronischer Bildsensor ( <i>„Charge Coupled Device“</i> ), Seite 25
CIR	<i>Color InfraRot</i> : Luftbild von einem Farbfilm mit Sensibilität für infrarote Strahlung mit 700–900 nm Wellenlänge (HILDEBRANDT 1996, Kapitel 4.1.5.4), Seite 11
$c$	Nebenbedingung bei der Optimierung, Constraint, Seite 138
D	Durchmesser eines Objektivs, Seite 32
D-GPS	Differenzielles GPS ( <i>„Differential Global Positioning System“</i> ), Seite 30
$\Delta\varphi$	Absolute Differenz von der Orientierung einer Trennlinie zu der Orientierung eines Pixels, Seite 104
$d$	Distanz, Seite 53
DOM	<i>Digitales OberflächenModell</i> , engl. DSM ( <i>„Digital Surface Model“</i> ), Seite 6
DR	Detektionsrate, Seite 91
DTM	<i>Digitales TerrainModell</i> , oder DGM, ( <i>Digitales GeländeModell</i> ), Seite 45
ENVI	Programmsystem zur Analyse und Visualisierung von Fernerkundungsdaten ( <i>„ENVironment for Visualizing Images“</i> ), Seite 19
$F$	Bildmatrix, Seite 23
FAR	Fehlalarmrate (falsch positiv), Seite 91
$f$	Brennweite, Seite 30
FoV	Seitl. Öffnungswinkel eines Sensors ( <i>„Field of View“</i> ), Seite 27
FRR	Fehlrückweisungsrate (falsch negativ), Seite 91

$F_V$	Binärbild, das die Konturlinie eines Gebäudes wiedergibt, Seite 130
$G$	Gauß-verteilter Faltungskern, Seite 130
GIS	Geografisches Informationssystem, Informationssystem, das der Bereitstellung von Fachinformationen unter Berücksichtigung ihres Raumbezuges dient, Seite 2
GPS	Globales Positionsbestimmungssystem („Global Positioning System“), Seite 30
$G$	Menge der Segmentrandpixel mit hoher Gradientenstärke, Seite 100
$h$	Häufigkeit (Histogramm), Seite 101
$h$	Höhe eines Punktes, Seite 48
$H$	Homogenitätsmaß, Seite 108
$\hat{H}$	Auf den Bereich $[0, 1]$ normiertes Homogenitätsmaß, Seite 109
HRSC	„High Resolution Stereo Camera“, Hochauflösende, digitale Stereokamera des DLR. Die Zusätze A, AX, AXW stehen für die luftgestützte Variante zur Erdfernerkundung („Airborne“), deren Fortentwicklung („eXtended“), sowie die Weitwinkelversion („Wide angle“), Seite 25
HSV	Farbraum, der aus dem Farbton, der Farbsättigung und der Helligkeit aufgespannt wird („Hue, Saturation, Value“); ähnlich sind der HSI- („Intensity“) und der gleichwertige HSL-Farbraum („Luminance“), Seite 76
IDL	Programmiersprache für die Analyse und Visualisierung von Bilddaten („Interactive Data Language“), Seite 19
INS	Inertiales Messsystem („Inertial Navigation System“), Seite 30
$\mathcal{J}$	Geordnete Menge an Indizes zu einem Vektor, Seite 108
$L$	Strahldichte vor einem Sensor, Seite 32
$l$	Länge einer Linie, Seite 104
$\mathbf{m}$	Merkmalsvektor, Seite 87
$m$	Steigung, Seite 53
$ \overline{\nabla} $	Anhand Gradientenorientierung gewichtet gemittelte Gradientenstärke, Seite 104
$ \nabla $	Gradientenstärke, Seite 23
$ \nabla ^{(loc)}$	Lokale Gradientenstärke (Unterdrückung der Gradientenstärke fern eines Gebäudepolygons), Seite 130
$N$	Nachbarschaftsfenster um aktuelles Pixel, Seite 47
nDOM	Auf Grundhöhe normiertes DOM, Seite 61

nDOM'	Vorläufiges durch den Abgleich mit umgebenden Höhenminima normiertes Oberflächenmodell, Seite 48
NDVI	Vegetationsindex basierend auf Bildinformation im roten und nah-infraroten Spektralbereich („Normalized Difference Vegetation Index“), Seite 70
NIR	Nahes InfraRot, Seite 70
$p$	Wahrscheinlichkeit bzw. Verteilungsfunktion dafür, Seite 87
PC	„Personal Computer“, Seite 19
$\phi$	Blickwinkel in Bezug zum Nadir, Seite 32
$\varphi$	Gradientenorientierung, Seite 23
$\hat{\varphi}$	Hauptgradientenorientierung, Seite 101
$q^{(T)}$	Gütekriterium zur Beurteilung von Trennlinien, Seite 112
$Q_{\text{rect}}^{(V)}$	Qualitätskriterium zur Beurteilung der Vertexkandidaten für die Rechtwinkligkeits-Nebenbedingung, Seite 149
$Q^{(V)}$	Aus der Gradientenstärke abgeleitetes Maß zur Beurteilung eines Polygons, Seite 133
$q^{(V)}$	Aus der Gradientenstärke abgeleitetes Maß zur Beurteilung eines Polygonsegments, Seite 133
$\mathcal{R}_{\text{all}}$	Die Menge der Polygonvertices, die einen nahezu rechten Winkel einschließen, Seite 139
$R_{\text{CCD}}$	Spektrale Empfindlichkeit eines Sensors, Seite 32
RGB	Farbraum, der aus den drei Farben Rot, Grün und Blau aufgespannt wird, Seite 72
$\rho$	Der von zwei benachbarten Polygonsegmenten eingeschlossene Winkel, Seite 139
$\varrho$	Orientierung eines Polygonsegments, Seite 133
$\bar{\varrho}$	Grundorientierung eines Gebäudepolygons, Seite 140
$\mathcal{R}_{\text{inner}}$	Polygonvertices aus $\mathcal{R}_{\text{all}}$ , deren Nachbarn ebenfalls in $\mathcal{R}_{\text{all}}$ enthalten sind, Seite 139
$\mathcal{R}'_{\text{inner}}$	Auswahl an Vertices aus $\mathcal{R}_{\text{inner}}$ , anhand derer die Hauptorientierung eines Polygons bestimmt wird, Seite 140
RSI	„Research Systems, Inc.“, Seite 19
S	Signal, Seite 32
$s_d$	DOM-Skalierungsfaktor bei der pyramidalen Verfeinerung der Geländehöhe, Seite 53
seg	Segment, Region, Seite 49
$\mathcal{S}_{\text{gi}}$	Segmentierung, Seite 48

$s$	Pixelgröße eines Bildes („Ground Sample Distance“) bzw. Abstand benachbarter CCD-Elemente, Seite 30
$s_w$	Optionaler Skalierungsfaktor für die Größe des Nachbarschaftsfensters bei der pyramidalen Verfeinerung der Geländehöhe, Seite 53
$\tau$	Transmission eines optischen Systems (Optik und Spektralfilter), Seite 32
$\theta$	Schwellwert, Seite 48
$\vartheta$	Lokale Orientierung der Skelettlinie eines Segments, Seite 101
$t_{\text{int}}$	Integrationszeit (Belichtungszeit eines CCD-Elements), Seite 29
$T$	Trennlinie, auch: alle an der Linie beteiligten Pixel, Seite 105
$T^{(g)}$	Binärer Vektor, der angibt, welches der Trennlinienpixel ein passendes Kantenpixel darstellt, Seite 108
$V$	Helligkeit über mehrere Bildkanäle („Value“), Seite 81
$v$	Fluggeschwindigkeit, Seite 29
$v$	Polygonstützstelle, Vertex (2-D), Seite 129
$\mathbf{v}$	Umsetzung der $n$ zweidimensionalen Polygonvertices in einen eindimensionalen Zustandsvektor der Länge $2n$ , Seite 141
$\mathcal{V}$	Menge der Bildpunkte, die von einem Polygonsegment abgedeckt werden, Seite 133
$V$	Geschlossener Polygonzug, Seite 129



# Literaturverzeichnis

- J. Albertz et al.** (1992). *The Camera Experiments HRSC and WAOSS on the Mars94 Mission*. In *International Archives of Photogrammetry and Remote Sensing*, Bd. 29, S. 216–227. Teil B1.
- A. A. Amini, S. Tehrani und T. E. Weymouth** (1988). *Using Dynamic Programming for Minimizing the Energy of Active Contours in the Presence of Hard Constraints*. In *Proceedings of the Second International Conference on Computer Vision*, S. 95–99.
- ASL (1999). *ASTER Spectral Library*. Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA. <http://speclib.jpl.nasa.gov/>.
- D. Ballard und C. Brown** (1982). *Computer Vision*. Prentice-Hall, Englewood Cliffs, NJ, USA.
- K. Barnard und G. Finlayson** (2000). *Shadow Identification using Colour Ratios*. In *Eighth Color Imaging Conference: Color Science, Systems and Applications*, S. 97–101. Imaging Science and Technology/SID.
- R. Bellman** (1957). *Dynamic Programming*. Princeton University Press.
- R. Bellman und R. Kalaba** (1965). *Dynamic Programming and Modern Control Theory*. Academic Press, New York.
- F. Bignone et al.** (Apr. 1996). *Automatic Extraction of Generic House Roofs from High Resolution Aerial Imagery*. In *European Conference on Computer Vision*. Cambridge, England.
- T. Brandtberg und F. Walter** (Okt. 1998). *Automated Delineation of Individual Tree Crowns in High Spatial Resolution Aerial Images by Multiple-Scale Analysis*. *Machine Vision and Applications*, Bd. 11(2), S. 64–73.
- C. Brechbühler** (1995). *Description and Analysis of 3-D Shapes by Parametrization of Closed Surfaces*. ETH No. 10979, ETH Zürich, Zürich.
- I. N. Bronstein und K. A. Semendjajew** (1991). *Taschenbuch der Mathematik*. Teubner Verlagsgesellschaft, Stuttgart, Leipzig, 25. Aufl.
- A. Brunn** (Febr. 1998). *Techniques for Automatic Building Extraction*. In *3rd Course in Digital Photogrammetry*, Kap. 11. Universität Bonn, Bonn.
- A. Brunn** (Apr. 2001). *Statistical Interpretation of DEM and Image Data for Building Extraction*. In *ASCONA 2001, International Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images* (Hg. **A. Gruen und M. Baltsavias**). Balkema, Niederlande.

- A. Brunn und U. Weidner** (1997). *Extracting Buildings from Digital Surface Models*. In *3D Reconstruction and Modelling of Topographic Objects*, Bd. 32, S. 27–34. International Archives of Photogrammetry and Remote Sensing. Teil 4-4W2.
- J. Burns, A. Hanson und E. Riseman** (Juli 1986). *Extracting Straight Lines*. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Bd. 8(4), S. 425–455.
- R. T. Collins et al.** (Apr. 1995). *Automatic Extraction of Buildings and Terrain from Aerial Images*. In *International Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images*. Ascona, Schweiz.
- R. T. Collins et al.** (Nov. 1998). *The Ascender System: Automated Site Modeling from Multiple Aerial Images*. *CVIU*, Bd. 72(2), S. 143–162.
- M. Cord, M. Jordan und J.-P. Cocquerez** (Mai 2001). *Accurate Building Structure Recovery from High Resolution Aerial Imagery*. *Computer Vision and Image Understanding*, Bd. 82(2), S. 138–173.
- T. H. Cormen, C. E. Leiserson und R. L. Rivest** (1990). *Introduction to Algorithms*. McGraw-Hill Book Company.
- G. Derksen, M. Peick und V. Ulbricht** (2001). *Die Automatisierte Liegenschaftskarte (ALK) des Landkreises Barnim. Vermessung Brandenburg*, (1), S. 31–41.
- O. Dissard et al.** (1997). *Above-Ground Objects in Urban Scenes from Medium Scale Aerial Imagery*. In GRÜN ET AL. (1997), S. 183–192.
- D. H. Douglas und T. K. Peucker** (1973). *Algorithms for the Reduction of the Number of Points Required to Represent a Line or its Caricature*. *The Canadian Cartographer*, Bd. 10(2), S. 112–122.
- R. O. Duda, P. E. Hart und D. G. Stork** (2001). *Pattern Classification*. John Wiley & Sons, Inc., 2. Aufl.
- W. Eckstein und O. Munkelt** (1995). *Extracting Objects from Digital Terrain Models*. In *Remote Sensing and Reconstruction for Three-Dimensional Objects and Scenes* (Hg. **T. Schenk**), Proc. of the SPIE 2572, S. 222–231.
- M. Ehlers, R. Janowsky und M. Gähler** (Jan. 2002). *New Remote Sensing Concepts for Environmental Monitoring*. In *Remote Sensing for Environmental Monitoring, GIS Applications and Geology* (Hg. **M. Ehlers**), Bd. 4545, S. 1–12. SPIE.
- C. Elvidge und R. Lyon** (1985). *Influence of Rock-Soil Spectral Variation on the Assessment of Green Biomass*. *Remote Sensing of Environment*, Bd. 17, S. 265–269.
- H. Feng und T. Pavlidis** (1975). *The Generation of Polygonal Outlines of Objects from Gray Level Pictures*. *IEEE Trans. Circuits and Systems*, Bd. 22, S. 427–439.
- D. Fére et al.** (1997). *On the Reconstruction of Urban House Roofs from Aerial Images*. In GRÜN ET AL. (1997), S. 87–96.

- A. Fischer et al.** (Nov. 1998). *Extracting Buildings from Aerial Images Using Hierarchical Aggregation in 2D and 3D*. *Computer Vision and Image Understanding*, Bd. 72(2), S. 185–203.
- R. Fletcher** (1987). *Practical Methods of Optimization*. John Wiley & Sons, Chichester, New York, Brisbane, Toronto, Singapur, 2. Aufl.
- J. D. Foley** (1990). *Computer Graphics: Principles and Practice*. Addison Wesley Publishing Company, Inc., 2 Aufl.
- S. Ford et al.** (1997). *Preliminary Results on the Analysis of HYDICE data for Information Fusion in Cartographic Feature Extraction*. In *SPIE Conference on Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision III*, Bd. 3072, S. 67–86.
- P. Fua und C. Brechbüler** (April 1996). *Imposing Hard Constraints on Deformable Models Through Optimization in Orthogonal Subspaces*. In *European Conference on Computer Vision*, S. 495–506.
- G. Funka-Lea** (1994). *The Visual Recognition of Shadows by an Active Observer*. Dissertation, University of Pennsylvania, Department of Computer and Information Science, GRASP Laboratories.
- G. Funka-Lea und R. Bajcsy** (Juni 1995). *Combining Color and Geometry for the Active, Visual Recognition of Shadows*. In *International Conference on Computer Vision*, S. 203–209. Cambridge, MA, USA.
- M. Gähler** (2000). *GIS-basierte Biotoptypenklassifikation auf Grundlage von multispektralen HRSC-A-Scannerdaten — Untersuchung der Einsatzmöglichkeiten am Beispiel der Elbinsel Pagensand*. Diplomarbeit, Universität Osnabrück, Fachbereich Kultur- und Geowissenschaften.
- M. Gerke, B. M. Straub und A. Koch** (2001). *Automatic Detection of Buildings and Trees from Aerial Imagery Using Different Levels of Abstraction*. *Publikationen der Deutschen Gesellschaft für Photogrammetrie und Fernerkundung*, (10), S. 273–280.
- S. Girard et al.** (Sept. 1998). *Building Detection from High Resolution Colour Images*. In *Intl. Symposium on Remote Sensing, EUROPTO'98*. SPIE, Bellingham, WA, USA.
- F. Gougeon** (1995). *A Crown-Following Approach to the Automatic Delineation of Individual Tree Crowns in High Spatial Resolution Aerial Images*. *Canadian Journal of Remote Sensing*, Bd. 21(3), S. 274–284.
- K. Greger** (1996). *Qualität von Geodaten — Anspruch und Wirklichkeit des Geodatenmarktes*. Techn. Ber., Runder Tisch GIS e.V. <http://www.rtg.bv.tum.de/index.php/article/articleview/79/1/24/>.
- A. Grün, E. P. Baltsavias und O. Henricsson** (Hg.) (1997). *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Bd. 2. ETH Zürich, Birkhäuser Verlag, Basel, Boston, Berlin.
- A. Grün und H. Dan** (1997). *TOBAGO — A Topology Builder for the Automated Generation of Building Models*. In GRÜN ET AL. (1997), S. 149–160.

- A. Grün, O. Kübler und P. Agouris** (Hg.) (1995). *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Bd. 1. ETH Zürich, Birkhäuser Verlag, Basel.
- E. Gülch** (Mai 1990). *Extraction of Contours in Digital Images by Active Contour Models*. In *International Archives of Photogrammetry and Remote Sensing*, Bd. 28 Teil 3, S. 211–220.
- E. Gülch, H. Müller und T. Läbe** (Sept. 1999). *Integration of Automatic Processes into Semi-Automatic Building Extraction*. In *Automatic Extraction of GIS Objects from Digital Imagery*, Bd. 32. International Archives of Photogrammetry and Remote Sensing, München. Teil 3-2W5.
- K. Gwinner** (1996). *Beeinflussung der Bildinformation in Fernerkundungsdaten durch Schatten*. In *Strahlung in Atmosphäre und Ozean* (Hg. **J. Fischer**), Bd. 4. Freie Universität Berlin, Institut für Weltraumwissenschaften.
- K. Gwinner und M. Schaale** (Juli 1997). *A Case Study on the Influences of Shadows and Shading on Multispectral Airborne Imaging Data*. In *Proc. of the Third International Conference on Airborne Remote Sensing*, S. 409–416. Kopenhagen, Dänemark.
- N. Haala und C. Brenner** (1997). *Interpretation of Urban Surface Models Using 2D Building Information*. In GRÜN ET AL. (1997), S. 213–222.
- N. Haala und C. Brenner** (Nov. 1998). *Interpretation of Urban Surface Models Using 2D Building Information*. *CVIU*, Bd. 72(2), S. 204–214.
- N. Haala und C. Brenner** (Juli 1999). *Extraction of Buildings and Trees in Urban Environments*. *ISPRS Journal of Photogrammetry and Remote Sensing*, Bd. 54(2–3), S. 130–137.
- R. Haralick, S. Sternberg und X. Zhuang** (1987). *Image Analysis Using Mathematical Morphology*. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Bd. 9(4), S. 523–550.
- E. Hauber et al.** (1996). *The High Resolution Stereo Camera (HRSC) for Mars96: Results and Outdoor Tests*. In *IAPRS96*, S. 349–354.
- A. J. Heller und L. H. Quam** (1997). *RADIUS: Image Understanding for Imagery Intelligence*, Kap. The RADIUS Common Development Environment. Morgan Kaufmann, San Mateo, CA, USA.
- M. Herman und T. Kanade** (1984). *The 3D MOSAIC Scene Understanding System: Incremental Reconstruction of 3D Scenes from Complex Images*. In *DARPA Image Understanding Workshop*, S. 137–148.
- G. Hildebrandt** (1996). *Fernerkundung und Luftbildmessung für Forstwirtschaft, Vegetationskartierung und Landschaftsökologie*. Herbert Wichmann Verlag, Heidelberg, 1. Aufl.
- A. Hoffmann** (Mai 2001). *Neue Ansätze zur Auswertung und Klassifizierung von sehr hochauflösenden Daten: Methoden der Segmentierung, der hierarchischen Klassifizierung und der per-Parcel-Methode mit Daten der Digitalen Kamera HRSC-A und ihre Anwendbarkeit für die Aktualisierung topographischer Karten*. Dissertation, Humboldt Universität, Berlin. <http://dochoost.rz.hu-berlin.de/dissertationen/hoffmann-andrea-2001-05-10/>.

- B. Horn und M. Brooks** (1989). *Shape from Shading*. MIT Press.
- T. Horprasert, H. D. und L. Davis** (Januar 2000). *A Robust Background Subtraction and Shadow Detection*. In *Proc. of the Asian Conference on Computer Vision*. Taipie, Taiwan.
- A. Huete, R. Jackson und D. Post** (1985). *Spectral Response of a Plant Canopy with Different Soil Backgrounds*. *Remote Sensing of Environment*, Bd. 17, S. 37–53.
- IAPRS96 (1996). *International Archives of Photogrammetry and Remote Sensing*, Bd. 31. Wien, Österreich.
- R. Irvin und D. McKeown** (1989). *Methods for Exploiting the Relationship Between Buildings and Their Shadows in Aerial Imagery*. *IEEE Trans. Systems, Man and Cybernetics*, Bd. 19(6), S. 1564–1575.
- ISPRS99 (September 1999). *Proc. of the ISPRS Workshop on Sensors and Mapping from Space*. ISPRS, Hannover. <http://www.ipi.uni-hannover.de/html/publikationen/1999/isprs-workshop>.
- C. O. Jaynes, A. Hanson und E. Riseman** (1997a). *Model-Based Surface Recovery of Buildings in Optical and Range Images*. In *Semantic Modeling for the Acquisition of Topographic Information from Images and Maps: SMATI97* (Hg. **W. Förstner und L. Plümer**), S. 211–227. Birkhäuser Verlag, Basel, Schweiz.
- C. O. Jaynes et al.** (Juni 1997b). *Building Reconstruction from Optical and Range Images*. In *IEEE Computer Vision and Pattern Recognition*, S. 380–386. SPIE.
- C. Jiang und M. O. Ward** (June 1992). *Shadow Identification*. In *IEEE Computer Vision and Pattern Recognition*, S. 606–612. Champaign, IL, USA.
- M. Kass, A. Witkin und D. Terzopoulos** (Januar 1988). *Snakes: Active Contour Models*. *International Journal of Computer Vision*, Bd. 1(4), S. 321–331.
- T. Kim und J.-P. Muller** (Sept. 1998). *A Technique for 3D Building Reconstruction*. *Photogrammetric Engineering and Remote Sensing*, Bd. 64(9), S. 923–930.
- Z. Kim, A. Huertas und R. Nevatia** (2000). *Automatic Description of Complex Buildings with Multiple Images*. In *Proc. 5th IEEE Workshop on Applications of Computer Vision*, S. 155–162.
- A. Koch** (Okt. 1999). *Analyse und Aufbereitung von Laser-Scanner-Aufnahmen*. Diplomarbeit, Universität Hannover, Institut für Photogrammetrie und GeoInformation, Hannover. <http://www.ipi.uni-hannover.de/html/lehre/diplomarbeiten/1999/andreas.pdf>.
- G. Konecny und J. Schiewe** (Aug. 1996). *Mapping from Digital Satellite Image Data with Special Reference to MOMS-02*. *ISPRS Journal of Photogrammetry and Remote Sensing*, Bd. 51(4), S. 173–181.
- K. Kraus und N. Pfeifer** (2001). *Advanced DTM Generation from LIDAR Data*. In *International Archives of Photogrammetry and Remote Sensing*, Bd. XXXIV. Annapolis, Maryland, USA.

- S. Krishnamachari und R. Chellapa** (Jan. 1996). *Delineating Buildings by Grouping Lines with MRFs*. *IEEE Trans. Image Processing*, Bd. 5(1), S. 164–168.
- T. Krüger** (Apr. 2002). *HRSC/ALK Berlin. Untersuchung der Passgenauigkeit von HRSC- und ALK-Daten*. Interner Bericht.
- T. Läbe und E. Gülch** (1998). *Robust Techniques for Estimating Parameters of 3D Building Primitives*. In *Proceedings of the ISPRS Symposium, Comm. II*. Cambridge.
- M. Larsen und M. Rudemo** (1997). *Estimation of Tree Positions from Aerial Photos*. In *Proceedings of the 1997 Swedish Symposium on Image Analysis* (Hg. **T. Lindeberg**), S. 130–134. Stockholm, Schweden.
- D. L. Lee und B. J. Schachter** (1980). *Two Algorithms for Constructing a Delaunay Triangulation*. *International Journal of Computer and Information Sciences*, Bd. 9(3), S. 219–242.
- T. M. Lillesand und R. W. Kiefer** (2000). *Remote Sensing and Image Interpretation*. John Wiley & Sons, Inc., New York, Chichester, Weinheim, Brisbane, Singapur, Toronto, 4 Aufl.
- C. Lin und R. Nevatia** (Nov. 1998). *Building Detection and Description from a Single Intensity Image*. *CVIU*, Bd. 72(2), S. 101–121.
- T. Lindeberg** (1990). *Scale-Space for Discrete Signals*. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Bd. 12(3), S. 234–254.
- Y.-T. Liow** (1991). *A contour tracing algorithm that preserves common boundaries between regions*. *Computer Vision Graphics and Image Processing*, Bd. 53(3), S. 313–321.
- Y.-T. Liow und T. Pavlidis** (Febr. 1990). *Use of Shadows for Extracting Buildings in Aerial Images*. *Computer Vision Graphics and Image Processing*, Bd. 49(2), S. 242–277.
- P. Lohmann und A. Koch** (September 1999). *Quality Assessment of Laser-Scanner-Data*. In ISPRS99. <http://www.ipi.uni-hannover.de/html/publikationen/1999/isprs-workshop/cd/pdf-papers/lohmman.pdf>.
- Q. Luong, P. Fua und Y. Leclerc** (Jan. 2002). *The Radiometry of Multiple Images*. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Bd. 24(1), S. 19–33.
- S. Marschner und D. Greenberg** (1997). *Inverse Lighting for Photography*. In *Proc. of the IST/SID Fifth Color Imaging Conference*, S. 262–265.
- S. Mason und E. Baltsavias** (1997). *Image-Based Reconstruction of Informal Settlements*. In GRÜN ET AL. (1997), S. 97–108.
- K. Masur** (1997). *Die Automatisierte Liegenschaftskarte (ALK)*. *Vermessung Brandenburg*, (1), S. 31–41.
- I. May** (2002). *Fortführung und Erweiterung von GDF (Geographic Data File) als Datengrundlage für Autonavigationssysteme*, Bd. 51 von *Berliner geographische Studien*. Univ.-Bibliothek der Techn. Univ. Berlin, Abt. Publ. Dissertation.

- H. Mayer** (Mai 1999). *Automatic Object Extraction from Aerial Imagery — A Survey Focusing on Buildings*. *Computer Vision and Image Understanding*, Bd. 74(2), S. 138–149.
- S. Mayer** (Oktober 1996). *Farbbasierte Objektlokalisierung mit einer Aktiven Kamera*. Studienarbeit, Universität Erlangen-Nürnberg, Erlangen.
- S. Mayer** (Sept. 2000a). *Extraction of Tree Groups from High-Resolution Digital Surface Models*. In *International Conference on Image Processing 2000*, Bd. 2, S. 712–715. IEEE, Vancouver, BC, Kanada.
- S. Mayer** (Juli 2000b). *Retrieving Object Borders and Height from High-Resolution Digital Surface Models*. In *World Multiconference on Systems, Cybernetics and Informatics 2000*, Bd. 5, S. 433–438. IIS, Orlando, FL, USA.
- S. Mayer** (Sept. 2002). *Gradient Based Splitting of Blocks of Houses in High-Resolution Ortho-Images*. In *International Conference on Image Processing 2001*, Bd. 1, S. 769–772. IEEE, Rochester, NY, USA.
- J. C. McGlone und J. A. Shufelt** (1994). *Projective and Object Space Geometry for Monocular Building Extraction*. In *IEEE Computer Vision and Pattern Recognition*, S. 54–61.
- D. M. McKeown** (1990). *Toward Automatic Cartographic Feature Extraction*. In *Mapping and Spatial Modeling for Navigation* (Hg. **L. Pau**), Bd. 65 von *NATO ASI Series F: Computer and Systems Sciences*, S. 149–180. Springer Verlag.
- R. B. McMaster** (Okt. 1987). *The Geometric Properties of Numerical Generalization*. *Geographical Analysis*, Bd. 19(4), S. 330–346.
- I. Mikic et al.** (2000). *Moving Shadow and Object Detection in Traffic Scenes*. In *International Conference on Pattern Recognition*, Bd. II, S. 321–324.
- M. Möller** (2001). *Monitoring of Urban Areas Using Digital Stereo Camera HRSC-A Imagery*. In *A Decade of Trans-European Remote Sensing Cooperation* (Hg. **Buchroithner**), S. 199–204. Balkema, Rotterdam, Niederlande.
- J.-P. Muller et al.** (1997). *Assessment of the Effects of Resolution on Automated DEM and Building Extraction*. In GRÜN ET AL. (1997), S. 233–242.
- G. Neukum und V. Tarnopolsky** (1990). *Planetary Mapping — The Mars cartographic Data Base and a Cooperative Camera Project for 1994*. *Geo-Information-Systeme*, Bd. 3(2), S. 20–29.
- G. Neukum et al.** (1995). *The Multiple Line Scanner Camera Experiment for the Russian Mars 96 Mission: Status Report and Prospect for the Future*. In *Photogrammetric Week '95* (Hg. **D. Fritsch und D. Hobbie**), S. 45–61. Wichmann, Heidelberg.
- G. Neukum et al.** (Nov. 2000). *Vom Mars zur Erde. DLR Nachrichten, Mitteilungen des Deutschen Zentrums für Luft- und Raumfahrt*, Bd. 99, S. 18–35.
- R. Nevatia, C. Lin und A. Huertas** (1997). *A System for Building Detection from Aerial Images*. In GRÜN ET AL. (1997), S. 77–86.

- F. O’Gorman und M. B. Clowes** (Apr. 1976). *Finding Picture Edges Through Collinearity of Feature Points*. *IEEE Trans. Computer*, Bd. 25(4), S. 449–455.
- T. Pavlidis** (1977). *Structural Pattern Recognition*. Springer Verlag, Berlin, Heidelberg, New York.
- T. Pavlidis** (1982). *Algorithms for Graphics and Image Processing*. Computer Science Press, Rockville, MD, USA.
- T. Pavlidis und Y.-T. Liow** (1990). *Integrating Region Growing and Edge Detection*. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Bd. 12(3), S. 225–233.
- A. Pinz** (1989). *Final Results of the Vision Expert System VES: Finding Trees in Aerial Photographs*. In *Wissensbasierte Mustererkennung* (Hg. **A. Pinz**), Bd. 49 von *OCG-Schriftenreihe*, S. 90–111. Oldenbourg.
- I. Pitas** (Februar 2000). *Digital Image Processing Algorithms and Applications*. John Wiley & Sons.
- R. Pollock** (Juni 1996). *The Automatic Recognition of Individual Trees in Aerial Images of Forests Based on a Synthetic Tree Crown Image Model*. Dissertation, University of British Columbia, Vancouver, BC, Kanada.
- D. Pouliot et al.** (Okt. 2002). *Automated Tree Crown Detection and Delineation in High-Resolution Digital Camera Imagery of Coniferous Forest Regeneration*. *Remote Sensing of Environment*, Bd. 82(2–3), S. 322–334.
- A. Prati et al.** (Dezember 2001). *Analysis and Detection of Shadows in Video Streams: A Comparative Evaluation*. In *IEEE Computer Vision and Pattern Recognition*, Bd. II, S. 571–576. Kauai, HI, USA.
- W. H. Press et al.** (1992). *Numerical Recipes - the Art of Numerical Computing, C Version*. Cambridge University Press, Cambridge, 2. Aufl.
- T. Ray** (1994). *A FAQ on Vegetation in Remote Sensing*. California Institute of Technology, Div. of Geological and Planetary Sciences, Pasadena, CA, USA. <ftp://kepler.gps.caltech.edu/pub/terrill/rsvegfaq.txt>.
- M. Ren, J. Yang und H. Sun** (Febr. 2002). *Tracing Boundary Contours in a Binary Image*. *Image and Vision Computing*, Bd. 20(2), S. 125–131.
- P. Rosin und T. Ellis** (1995). *Image difference threshold strategies and shadow detection*. In *Proceedings of the 6th British Machine Vision Conference*, S. 347–356. BMVA Press.
- J. W. Rouse et al.** (1973). *Monitoring Vegetation Systems in the Great Plains with ERTS*. In *Third ERTS Symposium, NASA SP-351*, Bd. 1, S. 309–317.
- M. Roux und D. M. McKeown** (1994). *Feature Matching for Building Extraction from Multiple Views*. In *IEEE Computer Vision and Pattern Recognition*, S. 46–53.
- H. Rüther, H. M. Martine und E. Mitalo** (Juli 2002). *Application of Snakes and Dynamic Programming Optimisation Technique in Modeling of Buildings in Informal Settlement Areas*. *ISPRS Journal of Photogrammetry and Remote Sensing*, Bd. 56(4), S. 269–282.



- M. Schaeffer** (Okt. 1999). *Filterung von Laser-Scanner-Aufnahmen*. Diplomarbeit, Universität Hannover, Institut für Photogrammetrie und GeoInformation, Hannover. <http://www.ipi.uni-hannover.de/html/lehre/diplomarbeiten/1999/schaeffer.pdf>.
- T. Schenk** (1999). *Digital Photogrammetry. Background, Fundamentals, Automatic Orientation Procedures*, Bd. 1. Terra Science, Laurelville, OH, USA.
- F. Scholten** (1996). *Automated Generation of Coloured Orthoimages and Image Mosaics Using HRSC and WAOSS Image Data of the Mars96 Mission*. In IAPRS96, S. 351–356. Teil B2.
- F. Scholten und F. Wewel** (2000). *Digital 3D-Data Acquisition with the High Resolution Stereo Camera - Airborne (HRSC-A)*. In *International Archives of Photogrammetry and Remote Sensing*, Bd. 33, S. 901–908. Amsterdam, Niederlande. Teil B4.
- F. Scholten, F. Wewel und S. Sujew** (2001). *Einsatz digitaler Kameras im Auflösungsbereich von 5 cm bis 1 m — Die neue HRSC-AX Kamerageneration in der geometrischen Analyse*. In *Photogrammetrie und Fernerkundung — Geoinformation: Geodaten schaffen Verbindungen* (Hg. **J. Albertz**), Bd. 11, S. 101–111. Publikationen der Deutschen Gesellschaft für Photogrammetrie und Fernerkundung, Berlin.
- K. Schutte** (1996). *Recognition of Buildings from Aerial Images*. In *Third Quinquennial Review 1991–1996* (Hg. **J. A. C. Bernsen et al.**), S. 211–225. Dutch Society for Pattern Recognition and Image Processing, Delft.
- J. A. Shufelt** (Juli 1996). *Projective Geometry and Photometry for Object Detection and Delineation*. Dissertation, Computer Science Dept., Carnegie Mellon Univ., Pittsburgh, PA, USA.
- J. A. Shufelt** (1997). *Geometric Constraints on Hypothesis Generation for Monocular Building Extraction*. In *Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision* (Hg. **D. M. McKeown, C. J. McClone und O. Jamet**), Bd. 3072 von *Proceedings of the SPIE*, S. 231–250. Teil III.
- A. R. Smith** (August 1978). *Color gamut transform pairs*. *ACM SIGraph, Computer Graphics*, Bd. 12(3), S. 12–19.
- G. Sohn und I. Dowman** (Juni 2001). *Extraction of Buildings from High Resolution Satellite Data*. In *International Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Bd. 2. ETH Zürich, Ascona, Schweiz.
- L. Spreeuwerts, K. Schutte und H. Zweitze** (Sept. 1997). *Modelbased Recognition and Parameter Estimation of Buildings from Multi-View Aerial Imagery Using Multi-Segmentation*. In *3D Reconstruction and Modeling of Topographic Objects*, Bd. 32. International Archives of Photogrammetry and Remote Sensing, Stuttgart. Teil 3-4W2.
- V. Steinhage** (1997). *On the Integration of Object Modeling in Automated Building Extraction from Aerial Images*. In GRÜN ET AL. (1997), S. 139–148.
- M. Stevens et al.** (1995). *Locating Shadows in Aerial Photographs Using Imprecise Elevation Data*. Techn. Ber. CS-95-105, Colorado State University, Computer Science Department, Fort Collins, CO, USA.

- B.-M. Straub und C. Heipke** (2001). *Automatic Extraction of Trees for 3-D-City Models from Images and Height Data*. In *Automatic Extraction of Man-Made Objects from Aerial and Space Images (III)* (Hg. **A. Gruen, E. P. Baltsavias und O. Henricsson**), Bd. 3, S. 149–160. ETH Zürich, Birkhäuser Verlag, Basel, Boston, Berlin.
- J. Toriwaki und T. Fukumura** (1978). *Extraction of Structural Information from Gray Pictures*. *Computer Vision Graphics and Image Processing*, Bd. 7, S. 30–51.
- C. J. Tucker** (1979). *Red and photographic linear combinations for monitoring vegetation*. *Remote Sensing of Environment*, Bd. 8, S. 127–150.
- C. Vestri und F. Devernay** (2001). *Using Robust Methods for Automatic Extraction of Buildings*. In *IEEE Computer Vision and Pattern Recognition*, I, S. 133–138.
- S. Vinson, L. Cohen und F. Perlant** (Jan. 2001). *Extraction of Rectangular Buildings using DEM and Orthoimage*. In *Scandinavian Conference on Image Analysis*. Bergen, Norwegen.
- G. Vosselman** (2000). *Slope Based Filtering of Laser Altimetry Data*. In *International Archives of Photogrammetry and Remote Sensing*, Bd. 33. Amsterdam, Niederlande.
- U. Weidner** (1997). *Digital Surface Models for Building Extraction*. In GRÜN ET AL. (1997), S. 193–202.
- U. Weidner und W. Förstner** (1995). *Towards Automatic Building Reconstruction from High Resolution Digital Elevation Models*. *ISPRS Journal of Photogrammetry and Remote Sensing*, Bd. 50(4), S. 38–49.
- F. Wewel** (1996). *Determination of Conjugate Points of Stereoscopic Three Line Scanner Data of Mars96 Mission*. In IAPRS96, S. 936–939. Teil B3.
- F. Wewel und M. Brand** (1999). *Geometrische Validierung des hochauflösenden multispektralen Mehrzeilenstereoscanners HRSC-A*. In *Photogrammetrie und Fernerkundung — Globale und Lokale Perspektiven* (Hg. **J. Albrecht**), Bd. 7, S. 245–252. Publikationen der Deutschen Gesellschaft für Photogrammetrie und Fernerkundung, Berlin.
- F. Wewel und F. Scholten** (2000). *Erstellung von Digitalen Oberflächenmodellen mit dem Mehrzeilen-Stereoscanner HRSC-A*. In *Photogrammetrie und Fernerkundung — Neue Anwendungen* (Hg. **J. Albrecht**), Bd. 8, S. 57–67. Publikationen der Deutschen Gesellschaft für Photogrammetrie und Fernerkundung, Berlin.
- F. Wewel, F. Scholten und K. Gwinner** (1999). *High Resolution Stereo Camera (HRSC) — Multispectral 3D-Data Acquisition and Photogrammetric Data Processing*. In *Proc. of the 4th Int. Airborne Remote Sensing Conference, Ottawa, Kanada*, Bd. 1, S. 263–272.
- F. Wewel et al.** (1998). *Digitale Luftbildaufnahme mit der HRSC-A — Ein Schritt in die Zukunft der Photogrammetrie*. *Photogrammetrie – Fernerkundung – Geoinformation*, Bd. 6, S. 337–348.
- G. Whitten** (Juli 1993). *Scale-Space Tracking and Deformable Sheet Models for Computational Vision*. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Bd. 15(7), S. 697–706.

- D. J. Williams und M. Shah** (Januar 1992). *A Fast Algorithm for Active Contours and Curvature Estimation*. *Computer Vision Graphics and Image Processing*, Bd. 55(1), S. 14–26.
- A. Woo, P. Poulin und A. Fournier** (Nov. 1990). *A Survey of Shadow Algorithms*. *IEEE Computer Graphics and Applications*, Bd. 10(6), S. 13–32.
- M. Wulder, K. Niemann und D. G. Goodenough** (Juli 2000). *Local Maximum Filtering for the Extraction of Tree Locations and Basal Area from High Spatial Resolution Imagery*. *Remote Sensing of Environment*, Bd. 73(1), S. 103–114.
- Y. Xia** (Okt. 1989). *Skeletonization via the Realization of the Fire Front's Propagation and Extinction in Binary Shapes*. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Bd. 11(10), S. 1076–1086.
- T. Yoon et al.** (September 1999). *Building Segmentation Using an Active Contour Model*. In ISPRS99. <http://www.ipi.uni-hannover.de/html/publikationen/1999/isprs-workshop/cd/pdf-papers/yoon.pdf>.
- A. Yuille, D. Cohen und P. Hallinan** (1989). *Feature Extraction from Faces Using Deformable Templates*. In *IEEE Computer Vision and Pattern Recognition*, S. 104–109.